3GPP TSG RAN WG1 #110 R1-220xxxx

Toulouse, France, August 22nd – 26th, 2022

Source: Moderator (OPPO)

Title: Summary#4 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. The company proposals are summarized, and offline proposals drafted passed on company contributions.

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. |

# Summary of Contributions and Offline Proposals

## Training and inference

### Training/inference at UE/NW side

In RAN1#109-e meeting, the following agreements were made:

|  |
| --- |
| 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 |

In this meeting, some contributions continue to discuss where the AI/ML model is trained and deployed. The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 1: AI/ML-based BM-Case 1, RAN1 studies further*   * *The AI/ML model is implemented with one-sided operation, i.e. training and inference are performed at the same side* * *For training, the UE can receive the RSRPs for a sparse Set B and as label the optimum beam ID from Set A* * *For inference, the AI/ML model can use the RSRSPs for a sparse Set B as input and infers the Top-K beams that will be used for final beam selection.*   *Proposal 2: AI/ML-based BM-Case 2, RAN1 studies further*   * *The AI/ML model is implemented with one-sided operation, i.e. training and inference are performed at the same side* * *For training, the AI/ML model can use as input N sets of RSRPs from N sparse Set B of historical information from the observation window and M optimum beam IDs as labels for the prediction window* * *For inference, the AI/ML model can use input N sets of RSRPs from N sparse Set B of historical information from the observation window and infers M sets of Top-K beams to be used for final beam selection in the prediction window*   *Proposal 3: For further study of BM-Case 1 and BM-Case 2,*   * *The same one-sided operation is supported, i.e. training and inference are performed at the same side for both BM-Case 1 and BM-Case 2* |
| Fujitsu[7] | *Proposal 1: Study spatial-domain DL beam prediction for mTRPs scenario.*   * *Both NW-side model and UE-side model should be studied.* |
| IDC[8] | *Proposal 1: Consider both AI/ML inference at NW side (Alt.1) and UE side (Alt.2) for both BM-Case1 and BM-Case2.* |
| Rakuten[10] | *Proposal 2: Single sided AI/ML (at the gNB side or the UE side) should be considered as baseline.* |
| CATT[13] | *Proposal 1: For the sub use case BM-Case1, consider both Alt.1 and Alt.2 for further study:*   * *Alt.1: AI/ML training at NW side;* * *Alt.2: AI/ML training at UE side.*   *Proposal 2: For the sub use case BM-Case1, consider following options for further study:*   * *Option1: AI/ML training and inference at NW side;* * *Option2: AI/ML training and inference at UE side;* * *Option3: AI/ML training at NW side and inference at UE side.*   *Proposal 7: For the sub use case BM-Case2, consider both Alt.1 and Alt.2 for further study:*   * *Alt.1: AI/ML training at NW side;* * *Alt.2: AI/ML training at UE side.*   *Proposal 8: For the sub use case BM-Case2, consider following options for further study:*   * *Option1: AI/ML training and inference at NW side;* * *Option2: AI/ML training and inference at UE side;* * *Option3: AI/ML training at NW side and inference at UE side.* |
| Intel[17] | *Observation 1: The ML model may reside either at UE or gNB* |
| Spreadtrum[18] | *Proposal 1: For both sub use cases BM-Case1 and BM-Case2, support AI/ML training at NW side.* |
| Charter[31] | *For BM-Case1 and BM-Case2, support AI/ML and deployment on NW side.* |

The first issue is where AI/ML model(s) is trained. In the last meeting, some related terminologies were agreed as working assumption:

|  |  |
| --- | --- |
| On-UE training | Online/offline training at the UE |
| On-network training | Online/offline training at the network |

Based on the tdocs submitted to this meeting, a small number of companies prefer to only consider On-UE training or On-network training. However, most companies seem to support both for this SI. Thus, moderator suggests to try the following proposal:

Proposal 2.1.1-1(H)

***Proposal 2.1.1-1: For the sub use case BM-Case1 and BM-Case2, support both Alt.1 and Alt.2 for AI/ML model training:***

* ***Alt.1: AI/ML model training at NW side (i.e., On-network training);***
* ***Alt.2: AI/ML model training at UE side (i.e., On-UE training).***
* ***Note: Whether it is online or offline training is a separate discussion.***

***Proposal 2.1.1-1a: 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 ~~(i.e., On-network training~~);***
* ***Alt.2: AI/ML model training at UE side ~~(i.e., On-UE training~~).***
* ***Note: Whether it is online or offline training is a separate discussion.***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Support |
| CATT | Support |
| ZTE | Support |
| Mediatek | Support |
| NEC | Support |
| Lenovo | Support |
| CAICT | Support |
| NVIDIA | Support |
| Futurewei | We are ok with the proposal. |
| Google | Support |
| Xiaomi | Since the definition of On-network training and On-UE training is not clear, we prefer to wait for the definition in framework agenda.  Mod: In AI 9.2.1, there is a definition for on-network/UE training. We understand that online/offline training used in the definition has not been defined so far. It seems not leading to some confusion on the understanding of on-network training and on-UE training. |
| Spreadtrum | Support |
| Panasonic | Support |
| vivo | We are okay to study both On-network training and On-UE training.  Just want to clarify what “support” means in this proposal. Does it mean this SI will consider both Alts for training discussion (to be included in the TR)? Or it means both will have spec impact in the future?  Mod: The former. Whether/how there is any spec impact will depend on further study. |
| Sony | Support both alternatives |
| OPPO | Support |
| Qualcomm | Support |
| Fujitsu | Support |
| HW/HiSi | Support |
| NTT DOCOMO | Support the proposal. |
| Ericsson | Support |
| Intel | Support |
| Apple | Support |
| Samsung | For the main bullet, not sure the meaning of ‘support’ here, probably the intention here is to support to further study. Also, the definition of ‘On-network training’ and ‘On-UE training’ is unclear to use. In our view, it is better to remove those terms. Please find our modification as below.  ***Proposal 2.1.1-1: For the sub use case BM-Case1 and BM-Case2, support both Alt.1 and Alt.2 for AI/ML model training for further study:***   * ***Alt.1: AI/ML model training at NW side ~~(i.e., On-network training)~~;*** * ***Alt.2: AI/ML model training at UE side ~~(i.e., On-UE training)~~.***   ***Note: Whether it is online or offline training is a separate discussion.***  Mod: updated |
| CMCC | Support |
| Mod | The proposal is updated according to the comments. Hope it is acceptable to all companies |
| vivo | OK for the updated proposal. |
| CAICT | Support the updated proposal. |
| Fujitsu | Support for the updated |
| Lenovo | Support the updated proposal. |
| Xiaomi | We are general fine with the updated proposal if most companies think it is unnecessary to mention the case of training on a non-3GPP server. |
| InterDigital | We think that other possible cases should be discussed as well as BM-Case1 and BM-Case2.  We prefer to discuss it together as shown in the below.  ***For ~~the sub use case BM-Case1 and BM-Case2~~ AI/ML for beam management, support both Alt.1 and Alt.2 for the study of AI/ML model training:***   * ***Alt.1: AI/ML model training at NW side ~~(i.e., On-network training~~);*** * ***Alt.2: AI/ML model training at UE side ~~(i.e., On-UE training~~).*** * ***Note: Whether it is online or offline training is a separate discussion.***   If companies want to focus on BM-Case 1 and 2, then we can consider the following update as a compromise.  ***At least f~~F~~or 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 ~~(i.e., On-network training~~);*** * ***Alt.2: AI/ML model training at UE side ~~(i.e., On-UE training~~).*** * ***Note: Whether it is online or offline training is a separate discussion.*** |
| CATT | Support FL’s updated proposal. |
| qualcomm | Support Proposal 2.1.1-1a. |
| HW/HiSi | Support the updated proposal  ***Proposal 2.1.1-1a: 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 ~~(i.e., On-network training~~);*** * ***Alt.2: AI/ML model training at UE side ~~(i.e., On-UE training~~).*** * ***Note: Whether it is online or offline training is a separate discussion.*** |
| Apple | FL’s update is fine. |
| MediaTek | Support FL’s proposal. |
| Ericsson | Support FL’s proposal. |
| Spreadtrum | Support the updated proposal. |
| Futurewei | Support the updates. |
| Charter | Support the updated proposal |

###### Proposal 2.1.1-1b (Closed)

According to inputs received so far, Proposal 2.1.1-1a seems accepted by almost companies. IDC suggested to keep the open for other sub use case (if some sub use case is agreed later). In moderator’s understanding, adding “at least” does not change anything of the proposal and hope it can be acceptable to all companies.

***Proposal 2.1.1-1b: At least for ~~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 ~~(i.e., On-network training~~);***
* ***Alt.2: AI/ML model training at UE side ~~(i.e., On-UE training~~).***
* ***Note: Whether it is online or offline training is a separate discussion.***

|  |  |
| --- | --- |
| Company | Comments |
| NTT DOCOMO | Support the proposal. |
| Mod | Agreement was made in online session |

Another issue is whether the AI/model training and inference are at the same node or different nodes. There would be four different alternatives:

* 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 NW side

By reviewing the tdocs, moderator got the impression that Alt. l and Alt.2 are supported or accepted by all companies, but there are some controversial views on Alt.3 and Alt.4. Thus, one possible way is to use the following proposal as a starting point for discussion. The proposal will be updated/refined according to the progress of discussion.

Proposal 2.1.1-2

***Proposal 2.1.1-2: 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:***

* ***Alt.1. AI/ML model training and inference at NW side***
* ***Alt.2. AI/ML model training and inference at UE side***
* ***Further discuss Alt.3 and Alt.4*** 
  + ***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 NW side***

***Proposal 2.1.1-2a: 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***
* ***Further discuss Alt.3 ~~and Alt.4~~*** 
  + ***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 NW side~~***

***Proposal 2.1.1-2b: 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***
* ***~~Further discuss Alt.3 and Alt.4~~*** 
  + ***~~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 NW side~~***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Support but prefer to deprioritize Alt3 and Alt4 |
| CATT | By our reviewing the company’s tdocs, we don’t think Alt.4 is supported by some companies. Moreover, for Alt.4, it is hard to merge or use a lot of different models sent from UE for NW. Thus, we prefer to delete Alt.4. The Proposal 2.1.1-2 can be updated as  ***Proposal 2.1.1-2a: 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:***   * ***Alt.1. AI/ML model training and inference at NW side*** * ***Alt.2. AI/ML model training and inference at UE side*** * ***Further discuss Alt.3 ~~and Alt.4~~***    + ***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 NW side~~*** |
| ZTE | For Alt.3, the case that AI/ML model inference at UE side is benefit for reducing reporting overhead and switching Rx beam in advance, and thus can be further studied. The study on Alt.4 can be deprioritized due to the absence of any clear advantages. |
| Mediatek | Support in principle, but prefer to focus on Alt 1 and 2. |
| NEC | We prefer to give high priority to Alt.1 and Alt.2. |
| Lenovo | We are fine with this proposal. For Alt3 and Alt4, we slightly prefer Alt.3 |
| CAICT | We also prefer Alt.3 is FFS and deprioritized Alt.4. |
| NVIDIA | Better to focus on Alt. 1 and Alt. 2. |
| Futurewei | We suggest deferring the decision whether to support Alt.3 and Alt.4 while proponents can present the corresponding benefits and corresponding overhead first. |
| Google | Support |
| Xiaomi | Wait for definition in framework agenda. |
| Spreadtrum | We prefer Alt 1. AI/ML model training has a severe test on the computing ability of the equipment. Compared with UE, gNB has powerful computing power, which is more conducive to complete model training efficiently. Alt3 can be considered if model transfer is supported in AI9.2.1 |
| Panasonic | We share similar view as ZTE and CATT. Alt 1 and Alt 2 can be prioritized and Alt 4 needs further justification. |
| vivo | We think all Alts deserve good study at this stage to understand each’s pros and cons in the whole LCM procedure. Without such study, we don’t know based on what principle we should prioritize or deprioritize anything.   1. Alt 1 and Alt 2 can be further studied as they are straight-forward for deployment. But they may need too much offline work in advance to switch or finetune a model, which causes latency if UE accesses into channel environment which hasn’t been pre-trained. 2. The benefit of Alt 3 or Alt 4 is it can enable dynamic switch of a model if it jointly work with level z. Take Alt 3 as an example. If a UE enters new channel environment, a new model can be used with a limited latency after the NW transfers the corresponding AI model to this UE.   The discussion is related with collaboration level discussion in 9.2.1. Alt 1 or Alt 2 may need the support of level-y, and Alt 3 or Alt 4 may need the support of level-z. We think it is better to wait for more progress in collaboration level discussion to see what is feasible from signaling perspective in Rel-18. |
| Sony | Federated learning can be considered as Alt. 4?  Mod: If I remember correctly, there is no contribution suggesting federated learning in BM. |
| OPPO | Support the FL proposal.  Since Alt.3 and Alt.4 are to be further discussed (c.f. Alt.1 & Alt.2 supported) in the proposal, Alt.3 and Alt.4 are somehow already deprioritized as asked by many companies. At early stage of SI, we are open to have each alternative (where to train and where to infer) studied and discussed. |
| Qualcomm | Support Alt 1, Alt 2, and Alt 3. For Alt 4, potential use cases should be elaborated to justify the alternative. |
| FUJITSU | Alt.1 and Alt.2 should have high priority. Alt4 can be deprioritized. |
| HW/HiSi | Agree with LGE, NEC, LENOVO, and PANASONIC, Alt 3 and Alt 4 would make it very complicated. |
| NTT DOCOMO | Support the proposal. |
| Ericsson | Alt 1, 2 ,3. We don’t see the feasibility of a large number of devices sending models to the NW. Remove alt 4. |
| Intel | OK to further study Alt. 1/2 and possibly 3. Alt. 4 does not seem very practical |
| Apple | We can study all the alternatives. |
| Samsung | Similar comments as above. Also, it seems that both Alt3 and Alt4 require model transfer between gNB and UEs. If this is the correct understanding, Alt3 and Alt4 are not preferred due to complexity and potential risk of AI model disclosure. Hence, we have the following wording suggestion.  ***Proposal 2.1.1-2: 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*** * ***~~Further discuss Alt.3 and Alt.4~~***    + ***~~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 NW side~~*** |
| CMCC | Support the proposal and also prefer to deprioritize Alt.4. |
| Mod | It seems almost companies agree to remove Alt.4. The views are still controversial on Alt.3. Thus, the proposal is updated. Let’s see whether we can achieve more progress |
| Qualcomm | Support Alt 1 and Alt 2. For one-sided AI/ML models, Alt 3 Alt 4, where training happens on behalf of the other side, does not make sense. **We oppose both Alt 3 and Alt 4**. |
| Mod2 | Based on the inputs, the following observations are made:  Alt.1 and Alt.2   * Support: all companies (Spreadtrum prefers Alt.1)   Alt.3   * Support: ZTE, vivo, Ericsson, Apple, * Not support or deprioritize: LGE, MTK, NEC, NVIDIA, FutureweI, Panasonic, OPPO, QC, FUJITSU, HW, Intel, Samsung,   Alt.4   * Support: vivo, Apple, * Not support or deprioritize: LGE, CATT, ZTE, MTK, NEC, Lenovo, CAICT, NVIDIA, FutureweI, Panasonic, OPPO, QC, FUJITSU, HW, Ericsson, Intel, Samsung, CMCC   As there are many companies not support Alt.3 and Alt.4, an updated version (2.1.1-2b) is provide to check whether it is acceptable to the proponents of Alt.3/4. |
| vivo | Not support. We don’t agree to deprioritize Alt 3 or Alt 4 at current stage. We think at least Alt 3 should be considered with same priority as Alt 1 and Alt 2. These two corresponds to different collaboration levels in 9.2.1. Hence all should be studied before there is any conclusion in 9.2.1. |
| CAICT | We prefer at lease Alt.3 should be considered. |
| Fujitsu | Alt3 and alt4 may involve the model transferring. If there are no big gains for alt.3 and alt.4, we’d better focus on alt.1 and alt.2 |
| Lenovo | We prefer to focus on Alt.1 and Alt.2 at this stage. |
| Xiaomi | We can focus on Alt 1 and Alt 2 first. |
| InterDigital | We are fine for prioritizing Alt. 1 and Alt. 2. |
| NEC | Support the updated proposal. |
| Samsung | We should focus on Alt 1 and Alt 2. |
| CATT | We don’t agree to delete Alt.3. At least Alt.3 can be further studied for collaboration level z defined in 9.2.1. |
| Qualcomm | Support updated proposal (Proposal 2.1.1-2b). As we are considering single-sided AI/ML models for the agreed use cases, it does not make sense to consider training at one side and inference at the other side. The nature of the inputs will be different at each side and Alt. 3 and Alt. 4 would inevitably involve model transfer and the model transfer comes with many intricacies and there’s a concern of proprietary AI/ML model disclosure as well and this would complicate the discussions. So given the timeline within Rel-18, suggest focusing on Alt. 1 and Alt. 2. |
| MediaTek | Support updated Proposal 2.1.1-2b. |
| Ericsson | Support FL’s updated proposal. |
| Spreadtrum | We think at least Alt 3 should be further studied according to the model transfer in 9.2.1. |
| Futurewei | We can focus on Alt.1 and Alt.2 in Rel-18. |
| Charter | Support Proposal 2.1.1-2b, especially Alt. 1 |

Proposal 2.1.1-2c

Based on the inputs received so far, the views are summarized as below:

**Alt.1 and Alt.2**

* Support: all companies (Spreadtrum, Charter prefers Alt.1)

**Alt.3**

* Support: ZTE, vivo, Ericsson, Apple, CATT, CAICT, Spreadtrum (6)
* Not support or deprioritize: LGE, MTK, NEC, NVIDIA, FutureweI, Panasonic, OPPO, QC, FUJITSU, HW, Intel, Samsung, Lenovo, xiaomi, IDC(15), Charter

**Alt.4**

* Support: vivo, Apple, (2)
* Not support or deprioritize: LGE, CATT, ZTE, MTK, NEC, Lenovo, CAICT, NVIDIA, FutureweI, Panasonic, OPPO, QC, FUJITSU, HW, Ericsson, Intel, Samsung, CMCC, Lenovo,xiaomi, IDC (21), Charter

Considering the current status, moderator suggest to deprioritize Alt.4 as the first step. Then, we focus on the controversial part of Alt.3.

***Proposal 2.1.1-2c: 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***
* ***[Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side]***
* ***~~Further discuss Alt.3 and Alt.4~~*** 
  + ***~~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 NW side~~***

|  |  |
| --- | --- |
| Company | Comments |
| NTT DOCOMO | Support the proposal. |
| NEC | Fine |
| vivo | We suggest to remove brackets for Alt 3. We don’t agree to deprioritize Alt 3 without any study. UE needs its specific model due to its own implementation and channel environment. It’s hard for UE to train and update its model due to high complexity. Hence we see the need of having Alt 3. |
| CATT | Share the same view as vivo. |
| PANASONIC | Support. |
| Xiaomi | Support the proposal |
| FUTUREWEI | Not sure what brackets means for Alt.3. If it is for FFS then please say it. We do prefer to leaving it FFS. |
| Fujitsu | support |
| Charter | We could support removing the brackets for Alt 3 |

###### Proposal 2.1.1-2d (H)

Based on the inputs, it seems most proponents of Alt.3 can accept FFS for Alt.3. Thus, Proposal 2.1.1-2d is provided to check the views

***Proposal 2.1.1-2d: 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***
* ***FFS: Alt.3. AI/ML model training at NW side, AI/ML model inference at UE side***

|  |  |
| --- | --- |
| Company | Comments |
| Futurewei | We support the proposal. |
| ZTE | We support the FL’s proposal. At current stage Alt.3 can be considered for further study. |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

### Online/offline training

There are discussions on the types of AI/ML model training for beam management. The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Observation 1: Given the dynamic nature in the propagation environment, online (reinforcement) learning may be a good alternative for AI/ML-based beam management in addition to offline learning approach like supervised learning.*  *Proposal 1: Study the standards impact, and pros and cons associated with both offline learning and online learning for AI/ML-based beam management.*  *Note: The definitions for offline learning and online learning are still being discussed under AI 9.2.1. The term “offline learning” in the proposal refers to supervised learning and “online learning” refers to reinforcement learning.* |
| Spreadtrum[18] | *Observation 1: Regarding AI/ML training for BM-Case1 and BM-Case2, offline training should be enough.* |
| Nokia[25] | *Proposal 10: Further study the BM-Case1 enhancements considering online/continual learning mechanisms.* |
| QC[27] | *Proposal 2: For training of UE-side AI/ML model, focus should be on offline training scenario, in which the development and training of the AI model for temporal beam prediction happens offline without the need to involve 3gpp signaling.*  *Proposal 7: For UE-side training, RAN1 should focus on offline training scenario for spatial domain beam prediction, in which the AI/ML model design and training does not involve 3gpp signalling.* |

Based on the tdocs submitted to this meeting and the inputs of the last meeting captured in FL summary [33], offline training can be supported by all companies. The controversial part is whether to support online training (i.e., reinforcement learning) or not:

* Some companies support online training, e.g., FUTUREWEI[1], Nokia[25]
* Some other companies prefer to only focus on offline training, e.g., Spreadtrum[18], QC[27]

Thus, Proposal 2.1.2 is suggested for the further discussion.

One thing should be noted that the terminologies of offline training and online training are still TBD in Agenda item 9.2.1.

###### Proposal 2.1.2

***Proposal 2.1.2: For the sub use case BM-Case1 and BM-Case2, support the following type(s) of AI/ML model training:***

* ***Alt.1. offline training***
* ***[Alt.2. online training e.g. for reinforcement learning]***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | This needs to be discussed after defining online and offline training in framework agenda |
| CATT | Agree with LGE. Prefer to wait the definition discussion in 9.2.1.  Moreover, we don’t think reinforcement learning is a kind of online training only. Reinforcement learning can also be a kind of offline training, e.g., reinforcement learning in an offline/virtual environment. |
| ZTE | Prefer to discuss the online training after it is clearly defined in the framework agenda. |
| MediaTek | Support Alt.1 only. Agree with ZTE that Alt. 2 needs further discussions. |
| NEC | Prefer to discuss it after defining online and offline training in 9.2.1. |
| Lenovo | We prefer to take Alt.2 as FFS. It can be discussed after we have clear understanding on online training. |
| CAICT | Alt.1 should be baseline and the definition of online and offline training should be further clarified in 9.2.1. |
| NVIDIA | Wait for further progress in 9.2.1. |
| Futurewei | We support the proposal. For AI/ML based beam management, online learning may have some benefits by continuously adapting to changes in the environment, particular with UE mobility (e.g., in BM-Case2). Whether to support online/offline can be left to vendor’s decision, the standards impact can be discussed together with the results when shared. |
| Google | Support |
| Xiaomi | Wait for definition in framework agenda. |
| Spreadtrum | Alt1 should be used as a starting point, and Alt2 can be discussed after the definition in 9.2.1 is stable |
| Panasonic | It is better to discuss this after progress in framework agenda 9.2.1. |
| vivo | We also agree that it can be discussed after we have a better understanding on what online or offline training means in 9.2.1. |
| Sony | Support both online and offline training |
| OPPO | For AI/ML model to infer beam(s), the baseline assumption can be the offline trained model. As for online training, as pointed by many companies, the definition on it seems lack of clarity by now.  We are fine with the FL proposal, by agreeing Alt.1 and keeping Alt.2 as in bracket or under FFS. |
| qualcomm | Agree with Alt. 1 and agree with LGE particularly for discussing Alt. 2. |
| Fujitsu | Prefer to discuss it after defining online and offline training. |
| HW/HiSi | We agree that the definition of online and off-line training needs to be clearly defined. If companies have different understanding here, it would not be fruitful to make a quick agreement here. We suggest to postpone. |
| NTT DOCOMO | Fine with the proposal. But we can defer this proposal until the definition of online/offline training is clear. |
| Ericsson | Support Alternative 1. Share the view that we should agree on the online/offline terminologies first. |
| Intel | Support Alt-1 at this point. Agree that definitions should be clarified before we agree on this. |
| Apple | On Alt. 1, we are open. On Alt. 2, we prefer any discussion if needed at all be taken to the general framework part. |
| Samsung | Same view as LGE. |
| CMCC | We prefer to prioritize offline training. |
| Mod | The views on different options are summarized as below:   * Defer the discussion related to online/offline training: LGE, CATT, NEC, NVIDIA, Xiaomi, Panasonic, vivo, * Alt.1: Offline training: ZTE, MTK, Lenovo, CAICT,FUTUREWEI, SPREADTRUM, * Alt.2: FUTUREWEI, GOOGLE, Sony,   According to the above information, the only feasible way is to discuss this issue later. |
| Lenovo | We are open to both online training and offline training. Companies can select what training method they would like to employ for evaluation. Whether to support offline training in spec can be further discussed when we have a common understanding on offline training. |
| InterDigital | We support Alt.1 and are open to Alt.2. We are fine for further studying Alt.2. |
| Qualcomm | We support Alt. 1 only. Alt 2 need concrete definition and understanding of online training coming from general framework (9.2.1). |
| Futurewei | We support Alt.1 and we also share the opinion from others that we can defer this proposal until the definition of online/offline training is clear. |
| Charter | We prefer a clear scope and definition for online training before adoption into the specification. |

## Details of BM-Case1 and BM-Case2 (except for input/output)

In RAN1#109e meeting, BM-Case1 and BM-Case2 were agreed for AI/ML-based beam management:

|  |
| --- |
| 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 |

Many contributions submitted to this meeting discuss more details of BM-Case1 and BM-Case2, e.g.,

* Input of AI/ML model
* Output of AI/ML model
* Construction of Set A and Set B and their relationship
* Scenario, Frequency ranges
* Generalization performance
* Other details

As the input/output of AI/ML model will be discussed in separate sections, this section will only discuss the remaining details (e.g., clarification of Set A and Set B).

### General views

There are some contributions discussing the high-level principle of AI/ML model inputs. The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| vivo[4] | *Proposal 10: Study two-step beam prediction scheme for improving generalization performance in BM-case1.*  *Proposal 12: Study two-step beam prediction scheme for improving generalization performance in BM-case2.*  *Proposal 16: Study the two possible AI-based beam prediction solutions, i.e. beam pair prediction scheme and two-step beam prediction scheme, and its specification impact, both considering generalization aspects like Set B construction, supported number of Tx/Rx beams, various number of antenna configurations, etc.* |
| IDC[8] | *Proposal 4: 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.*  *Proposal 5: Companies supporting the alternative should provide more details for predicting L1-RSRP values without any beam information.* |
| Nokia[25] | *Proposal 19: For BM-Case2 temporal domain beam prediction, RAN1 should study the impact of the historical data length as well as on accuracy for the prediction future steps.* |

Each proposal in the above table is only discussed in one tdoc. The proponent(s) is encouraged to discuss with other companies and get more supporters.

In the tdocs, different companies have different assumptions for BM-Case1 and BM-Case2 regarding what beam(s) is predicted. In general, three different assumptions were discussed in the tdocs or used in the evaluations:

* Tx beam
* Rx beam
* A pair of Tx beam and Rx beam (beam pair)

To facilitate the discussion and make the evaluation results comparable, it is beneficial to make it clear. Thus, a proposal is suggested as below for further discussion:

Proposal 2.2.1 (H)

***Proposal 2.2.1: For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams with potential down-selection:***

* ***Alt.1: Tx beam prediction***
* ***Alt.2: Rx beam prediction***
* ***Alt.3: Beam pair prediction (a beam pair consists of a Tx beam and a corresponding Rx beam)***

***Proposal 2.2.1a: For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams with potential down-selection:***

* ***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 not have spec impact***
* ***Note2: Rx beam is part of UE implementation and how/which Rx beam is used is transparent to the spec***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Ambiguous whether Rx beam prediction has any specification impact. |
| CATT | We think finally the aim is to predict a best beam pair, not only Tx beam or Rx beam in reality. Thus, we prefer Alt.3. |
| ZTE | Both Alt.1 and Alt.3 can be considered. We think the difference of Alt.1 and Alt.3 is that whether measured RSRPs of all Rx beams (or a pre-determined Rx beam) or measured RSRPs of partial sampled Rx beams are input to the AI model. |
| mediatek | We prefer Alt. 1 and Alt. 3. |
| NEC | We prefer to study Alt.1 and Alt.3. |
| Lenovo | From spec point of view, it seems that Rx beam prediction dost not have any spec impact. So we support Alt.1 since we understand AI 9.2.3.2 should focus on the spec supporting on AI BM. |
| CAICT | Alt. 3 is preferred. |
| NVIDIA | In terms of priority, Alt. 1 > Alt. 3 > Alt. 2 |
| Futurewei | We support Alt.3 and we agree with LGE that Rx beam prediction may be done at UE without standards impact. |
| Google | We suggest changing “Tx” into “NW” and “Rx” into “UE” to avoid potential misunderstanding for UL/DL BM.  In our view, UE beam should be transparent and spec impact related study should focus on Alt1. |
| XIaomi | From our understanding, the difference between Alt 3 and Alt.1/Alt 2 is the joint Tx and Rx beam prediction or separate Tx/Rx beam prediction. And joint Tx and Rx beam prediction means both Tx beam and Rx beam are predicted with one AI model/procedure, like beam management procedure P1. Separate Tx/Rx beam prediction means Tx beam and Rx beam is predicted with different AI model/ procedure, like beam management procedure P2 and P3. Thus we suggest to revise the proposal as below:  ***Proposal 2.2.1: For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams with potential down-selection:***   * ***Alt.1: separate Tx beam and/or Rx beam prediction*** * ***Alt.2: joint Tx and Rx beam predication, i.e., Beam pair prediction (a beam pair consists of a Tx beam and a corresponding Rx beam)*** |
| Spreadtrum | We prefer Alt. 3. |
| Panasonic | Similar concern as Google, but we suggest to change “Tx” into “DL Tx”, and “Rx” into “DL Rx” for better clarity.  We support to priority Alt 1. |
| vivo | We are okay to further study this aspect.  We think to get a best beam pair to use, we can use Alt 3 independently, or we use Alt 1 and Alt 2 jointly based on a two-step manner.  Note that in the current specification, we only have Alt 1 (P2) or Alt 2 (P3) configuration, but we don’t have a dedicated configuration to support Alt 3 (P1). Hence to support Alt 3, a new CSI-RS configuration to support P1 is needed. |
| Sony | Support both Tx beam prediction and Rx beam prediction based on the use cases presented in this meeting |
| OPPO | Since BM-Case1 and BM-Case2 are defined in RAN1#109e as for DL beam prediction in spatial and temporal domain respectively, we understand these alternatives are also targeted for DL only.  We also believe that Alt.1 and Alt.3 are more aligned with NR beam management framework and Alt.2 needs more discussion on its use case. |
| qualcomm | Support Alt. 1 and Alt. 3, as the motivation for supporting them is clear for both UE-side and gNB-side AI/ML model use cases for beam prediction. For Alt. 2, we agree with LGE. |
| Fujitsu | Both alt.1 and alt.3 are considered. But for alt.1, it needs to be clarified what’s assumption about RX beam (e.g. wide RX beam, all RX beams or any pre-defined RX beam) |
| HW/HiSi | We are not sure if Rx beam selection has any specification impact, considering the Rx beams are just implementation related. Therefore, we prefer Alt.1, or at least Alt.1 with higher priority. |
| NTT DOCOMO | Support the proposal. Even Rx beam prediction might have potential specification impacts such as assistance information for better beam prediction. |
| Ericsson | We support all three alternatives |
| Intel | We support all alternatives. Ok to change Tx and Rx to NW and UE beam as suggested by Google to avoid further ambiguity. |
| Apple | We don’t see why Rx team is mentioned at all, as that should be part of UE implementation. So Alt. 1 is fine, but we have concern on the rest. |
| Samsung | We are fine with the proposal to facilitate the further discussion.  We think alt.1 should be the basic case. |
| CMCC | We prefer Alt.3. |
| Mod | The proposal is updated   * Most companies thought Rx beam prediction have no spec impact while some other company think there may have some spec impact. There are some evaluations with the assumption of Rx beam prediction. Considering this is the initial study phase of AI, Mod suggest to include this case and the group can decide whether this has spec impact or not later. Note 1 is added and hope it can the concerns of some companies * The alternative involving “Rx beam” is to facilitate the discussion. Note2 is added and hope it can address some concerns |
| vivo | OK with the proposal. |
| CAICT | Fine with the update. |
| Fujitsu | Support for the updated |
| Lenovo | We understand that AI9.2.3.2 are focus on the spec impact of AI for BM, and if Alt2 can be supported without spec impact, we are confused why it should be included for down-selection. |
| Xiaomi | Ok with the proposal. As for Alt 2, we think the RS overhead can be reduced based on some signaling exchanged between UE and gNB. |
| InterDigital | We believe that these alternatives are a good basis for further study and support the proposal. |
| NEC | Support the updated proposal. |
| Samsung | In our view, if we would like to discuss the specification impact, it is better to discuss it with the combination of UE/gNB side AI model inference. For example, if gNB side AI inference is adopted, both Alt.2 and Alt.3 may have specification impact. To be safe, we suggest the following wording.  ***Proposal 2.2.1a: For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams with potential down-selection:***   * ***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 not or may not have spec impact*** * ***~~Note2: Rx beam is part of UE implementation and how/which Rx beam is used is transparent to the spec~~*** |
| CATT | OK for the updated and also fine with Samsung’s suggestion. |
| Qualcomm | The feasibility of the alternatives clearly depend on which side the AI/ML model inference takes place. We suggest the following clarification and agree with Samsung’s suggestion. So we suggest replacing previous ‘Note2’ with the following one:   * ***Note2: The feasibility and down-selection may be different depending on whether the inference is at UE side or at gNB side.*** |
| HW/***HiSi*** | Thank you for the update. We think the updated proposal looks better now. We would prefer Alt 1, but would not object if this is the majority view.  ***Proposal 2.2.1a: For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams with potential down-selection:***   * ***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 not have spec impact*** * ***Note2: Rx beam is part of UE implementation and how/which Rx beam is used is transparent to the spec*** |
| Apple | We support Alt. 1. Think about generalization, how to build a Rx beam predictor for UE vendor A, B, C, D..? Rx beam design is UE implementation choice, that does not fall under 3GPP specification.  Mod: This is study item. We can make some conclusion in future there is no spec impact on some alternative(s). |
| Ericsson | Support FL’s updated proposal. |
| Spreadtrum | Support the updated proposal. |
| Futurewei | We are ok with the updated proposal. |
| Charter | Support the updated proposal, especially Alt 1 |

###### Proposal 2.2.b (closed)

Based on the inputs received so far, majority companies can accept the main content although some companies have preference on some specific alternatives.

For the notes, there is some suggestions. The proposal is updated by combining the suggestions from SS and QC and some monitor modification on top of their proposed texts.

***Proposal 2.2.1b: For the sub use case BM-Case1 and BM-Case2, further study the following alternatives for the predicted beams with potential down-selection:***

* ***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***
* ***~~Note2: Rx beam is part of UE implementation and how/which Rx beam is used is transparent to the spec~~***
* ***Note2: The feasibility of Rx beam and down-selection of alternatives may be different depending on whether the inference is at UE side or at gNB side.***

|  |  |
| --- | --- |
| Company | Comments |
| NTT DOCOMO | Support the proposal. |
| Mod | Agreement was made in online session |

### Construction of Set A and Set B

In RAN1#109e meeting, some alternatives for constructions of Set A/B were agreed for BM-Case1 and BM-Case2 as below:

|  |
| --- |
| 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  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 |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 4: For BM-Case 1, for the definition of Set B and Set A, both Alt.1 and Alt.2 can be considered but detailed analysis and comparisons should be provided.*   * *Alt.1: Set B is a subset of Set A* * *Alt.2: Set A and Set B are different*   *Proposal 5: For BM-Case 2, to provide sufficient flexibility for the AI/ML design, the selection of Set B can be:*   * *Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)* * *Alt.2: Set B is a subset of Set A (Set A and Set B are not the same)* |
| TCL[3] | *Proposal 3: The subsets of beams at the gNB side and UE side, can be constructed with the assistance of an ML model to reduce the beam training overhead.* |
| Vivo[4] | *Proposal 15: Slightly prefer Alt.1, i.e. Set B is a subset of Set A, as representative sub use case for further study in both BM-case1 and BM-case2, due to lower simulation complexity, but we can live with other alternatives.* |
| ZTE[5] | *Proposal 2: The sub-sampling based method in Alt.1 can serve as a starting point for the study of spatial domain beam prediction.*  *Proposal 3: The association in reference signals between two sets with different beam widths need to be further studied.*  *Proposal 6: Regarding the beam set construction, Alt.3 can be used as a benchmark, while Alt.1 and Alt.2 are deferred until the evaluation of the spatial domain beam prediction in BM-Case1 has achieved sufficient progress.* |
| IDC[8] | *Proposal 2: 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 3: 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.* |
| Rakuten[10] | *Proposal 1: Both of the following use cases should be considered for the AI/ML based beam management framework: “Set B is a subset of Set A”, and “Set A and Set B are different”.* |
| OPPO[11] | *Proposal 2: For BM-Case1, Set B can be a subset of Set A with fixed pattern.*  *Proposal 6: For BM-Case2, Set B and Set A can be the same.* |
| CATT[13] | *Proposal 4: For the Alt.2 of sub use case BM-Case1, i.e., Set A and Set B are different, some relationship is needed between beams in Set A and Set B.*   * *For example, the beams in Set A and Set B cover the similar area.*   *Proposal 9: For the sub use case BM-Case2, all of the following alternatives can be further studied:*   * *Alt.1: Set A and Set B are different;* * *Alt.2: Set B is a subset of Set A;* * *Alt.3: Set A and Set B are the same.* * *Note: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.* |
| NEC[14] | *Proposal 1: For BM-Case1, support the following alternatives for further study:*   * *Alt.1: Set B is a subset of Set A.* * *Alt.2: Set B and Set A are different.*   *Proposal 5: For BM-Case2, support the following alternatives for further study:*   * *Alt.1: Set A and Set B are different.* * *Alt.2: Set B is a subset of Set A (Set A and Set B are not same).* * *Alt.3: Set A and Set B are the same.* |
| Lenovo[15] | 1. *The number of beams within the prediction beam set, i.e., beam Set A is less than the number of beams within the measurement beam set, i.e., beam Set B.* |
| Spreadtrum[18] | *Proposal 2: For Alt.1 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.* |
| Xiaomi[19] | *Proposal 3: For spatial domain beam prediction, consider set B is a subset of set A with high priority.* |
| CAICT[20] | *Proposal 3: For spatial-domain beam prediction at UE side, Set B should be a subset of Set A. Set B is randomly chosen as baseline.*  *Proposal 4: For spatial-domain beam prediction at gNB side, the correspondence of Set B and Set A could be flexible.*  *Proposal 5: For time-domain beam prediction, Set A and Set B could be considered as the same.* |
| Samsung[21] | *Proposal 1: For the sub use case BM-Case1, consider to define Set C for AI/ML inference at NW side.*   * *Set C consists of the beams reported by UE from Set B.*   *Proposal 4: For the sub use case BM-Case2, consider to define Set C for AI/ML inference at NW side.*   * *Set C consists of the beams reported by UE from Set B.* |
| LGE[22] | *Proposal #1: For the relation between Set A and Set B of BM-Case1, both Alt1 and Alt2 can be considered for this SI and potential subsequent WI in Rel-19, and which Alt to apply could be up to NW’s implementation choice.*  *Proposal #4: For the relation between Set A and Set B of BM-Case2, consider Alt3 as a baseline to see performance of TD prediction and SD prediction separately.* |
| Ericsson[24] | 1. *Avoid restricting the beam configuration by using wide and narrow beam terminology when defining alternatives for beam set A and B.* 2. *Avoid restricting beam configuration alternatives at this stage by defining QCL relations between set A and B* |
| Nokia[25] | *Proposal 12: For DL Tx beam prediction Set B is different to Set A, consider Set B is a wide beam codebook and Set A is a refined beam codebook.*  *Proposal 13: For Set B is different to Set A, the Set B wide beam measurements can come from the measurements from SSB and/or CSI-RS.*  *Proposal 15: For BM-Case1 with Set A/B consider 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.*  *Proposal 21: 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.* |
| MTK[26] | *Proposal 1: RAN1 should discuss and agree on the relationship between Set-A and Set-B.*   * *Alt-1: Set-B is a subset of Set-A.* * *Alt-2: Set-B is different type from Set-A.* * *Both Alt-1 and Alt-2.*   *Proposal 2: Discussions are needed on how to determine Set B from Set-A.*  *Proposal 4: Agreements are needed on how to determine Set B from Set-A, if, and when the two sets are different.* |
| Panasonic[30] | *Observation 1: No need to down-select between Alt.1 (Set B is a subset of Set A) and Alt.2 (Set A and Set B are different) for BM-Case 1.*  *Observation 2: Alt 3 (Set A and Set B are the same) can be prioritized for the study of BM-Case 2.* |

The views of tdocs are summarized in the following tables:

|  |  |
| --- | --- |
| BM-Case 1 | |
| Set A and Set B are different(e.g. Set A consists of narrow beams and Set B consists of wide beams) | Huawei[2] , IDC[8], Rakuten[10], CATT[13], NEC[14], Spreadtrum[18], LGE[22], Nokia[25], Panasonic[30] |
| Set B is a subset of Set A | Huawei[2], vivo[4], ZTE[5], IDC[8], Rakuten[10], OPPO[11], CATT[13], NEC[14], Spreadtrum[18], Xiaomi[19], CAICT[20], LGE[22], Panasonic[30], Charter[31]. |

|  |  |
| --- | --- |
| BM-Case 2 | |
| Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams) | Huawei[2] , IDC[8], Rakuten[10], CATT[13], NEC[14], |
| Set B is a subset of Set A (Set A and Set B are not the same) | Huawei[2], vivo[4], IDC[8], Rakuten[10], CATT[13], NEC[14], Charter[31] |
| Set A and Set B are the same | ZTE[5], OPPO[11], CATT[13], NEC[14], Spreadtrum[18], CAICT[20], LGE[22], Nokia[25], Panasonic[30] |

From the above 2 tables, we can see that each alternative of BM-Case1 and BM-Case2 has a considerable number of supporting companies. It seems difficult for the group to down-select or prioritize some alternatives over the other ones. Meanwhile, tdocs showed the meaningful use case(s) for each alternatives. Thus, one possible way is to support all alternatives for the SI.

###### Proposal 2.2.2-1 (Closed)

***Proposal 2.2.2-1: For the sub use case BM-Case1, support the following alternatives:***

* ***Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)***
* ***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 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.***

***Proposal 2.2.2-1a: For the sub use case BM-Case1, support the following alternatives for further study:***

* ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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 narrow and wide beam terminology is for SI discussion only and have no specification impact~~***
* ***Note2~~3~~: The ~~codebook constructions~~ beam patterns of Set A and Set B can be clarified by the companies.***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Support |
| CATT | Support |
| ZTE | Support. Both alternatives can be studied. In Alt.1, the beam sets for measurement and for prediction originate from the same codebook with the same beam width, which is relatively simple to be implemented. Alt.2 may matches more with the current spec that SSB (wide beams) have to be sent anyway. Nevertheless, for Alt.2, how to generated the wide beams and the associated impact on the quality of existing communication services (cell coverage, etc.) need to be further studied. |
| Mediatek | Support both Alt. 1 and Alt. 2. |
| NEC | Support |
| Lenovo | Support |
| CAICT | Support |
| NVIDIA | Support |
| FUTUREWEI | We support the proposal. |
| Google | Suggest one minor revision for the last note as follows with regard to the case for non-codebook based BM.   * ***Note3: The beam patterns of Set A and Set B can be clarified by the companies.***   Mod: updated |
| Xiaomi | Support |
| Spreadtrum | Support |
| PANASONIC | We support to study both. |
| vivo | We are okay to further study these two Alts in this SI. Like our previous comment, what “support” means in this BM study should be clarified.  Mod: updated |
| OPPO | Support |
| Qualcomm | Support |
| Fujitsu | Support. |
| HW/HiSi | Support |
| NTT DOCOMO | Support the proposal. |
| Ericsson | Support. Except that we don’t see a need for including the notes on wide and narrow beam. We propose to remove such terms.  Mod: updated |
| Intel | Support |
| Apple | support |
| Samsung | Similar as the comments above, we have the wording suggestion as below.  ***Proposal 2.2.2-1: For the sub use case BM-Case1, support the following alternatives for further study:***   * ***Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)*** * ***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 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.***   Mod: updated |
| CMCC | Support |
| Mod | A new version is provided based on the inputs |
| vivo | OK with the proposal |
| CAICT | Support |
| Fujitsu | support |
| Lenovo | Support Proposal 2.2.2-1a. |
| InterDigital | Support |
| NEC | Support the updated proposal. |
| Samsung | Ok with the updated proposal. |
| CATT | Support |
| Qualcomm | Support Proposal 2.2.2-1a. |
| HW/HiSi | Support the updated proposal:  ***Proposal 2.2.2-1a: For the sub use case BM-Case1, support the following alternatives for further study:***   * ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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 narrow and wide beam terminology is for SI discussion only and have no specification impact~~*** * ***Note2~~3~~: The ~~codebook constructions~~ beam patterns of Set A and Set B can be clarified by the companies.*** |
| Apple | support |
| MediaTek | Support |
| Ericsson | Support |
| Spreadtrum | Support the updated proposal. |
| Mod | Please see the section of “Summary of discussion”.  The discussion is closed |

Proposal 2.2.2-2 (H)

***Proposal 2.2.2-2: For the sub use case BM-Case2, support the following alternatives:***

* ***Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)***
* ***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: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.***
* ***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.***

***Proposal 2.2.2-2: For the sub use case BM-Case2, support the following alternatives for further study:***

* ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.***
* ***~~Note2: The narrow and wide beam terminology is for SI discussion only and have no specification impact~~***
* ***Note2~~3~~: The ~~codebook constructions~~ beam pattern of Set A and Set B can be clarified by the companies.***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Alt1 and Alt2 will combine and mix TD and SD prediction aspects together so it will make hard to draw observations/conclusions from each aspect. Thus for BM-Case2, we suggest focusing on Alt3. Of course, Alt1 and Alt2 can be treated after sufficient progress of SD and TD prediction, e.g. in later phase of SI or in WI phase. |
| CATT | Alt.3 only focuses on temporal DL beam prediction. However, if AI/ML model inference is at NW side, the Alt.3 needs UE to feedback historic measurement results among all of Tx and Rx beam pairs. The reporting overhead is very large. If AI/ML model inference is at UE side, gNB needs to configure UE to measure all of Tx and Rx beam pairs in historic measurement, which will also cause large RS consumption. Alt.1 and Alt.2 are more realizable method for temporal DL beam prediction in reality. Thus, we are open for all alternatives for BM-Case2, i.e., we support the current Proposal 2.2.2-2. |
| ZTE | Support the FL’s proposal. Consider that Alt.1 and Alt.2 are actually a hybrid method of spatial domain beam prediction and temporal beam prediction. In other word, if the beam set for measurement and the beam set for prediction are different, the spatial domain beam prediction algorithm may be an essential precondition for the temporal beam prediction study in Alt.1 and Alt.2. Thus, Alt.3 can be used as a benchmark, while Alt.1 and Alt.2 are deferred until the evaluation of the spatial domain beam prediction in BM-Case1 has achieved sufficient progress. |
| mediatek | We prefer Alt. 3 to be the baseline. Alt.1 and Alt. 2 can also be studied, as these two are hybrid of Spatial and Temporal beam prediction. |
| NEC | Alt.3 involves only time-domain related prediction, while for Alt.1 and Alt.2, in addition to time-domain related prediction, it also involves spatial-domain related prediction. From the perspective of “simple to complex”, we prefer to study Alt.3 firstly. Both Alt.1 and Alt.2 can be studied later. |
| Lenovo | We support to focus Alt.3 on BM-Case2 while Alt.1 and Alt.2 are benefit to for BM-Case1 because of the reduced RS overhead. |
| CAICT | Alt.3 should be baseline. |
| NVIDIA | Better to focus on Alt. 3. |
| Futurewei | We are ok with the proposal. |
| Google | We do not think Alt3 should be the baseline. The benefit for Alt3 is minor.  Suggest one minor revision for the last note as follows with regard to the case for non-codebook based BM.   * ***Note3: The beam patterns of Set A and Set B can be clarified by the companies.*** |
| Xiaomi | We prefer Alt 3 as the baseline for BM case 2, since Alt 1 and Alt 2 are the combination of BM case 1 and BM case 2. |
| Spreadtrum | Alt.1 and Alt.2 are used to predict the beam in spatial domain. If temporal domain prediction and spatial domain prediction are conducted together, it will be difficult to analysis which part of the performance loss is cause by temporal domain prediction. So we prefer to study Alt.3 with high priority. |
| PANASONIC | Alt 3 can be prioritized in order to obtain clear insights on temporal and spatial beam prediction individually. |
| vivo | We are okay to further study these two Alts in this SI. Like our previous comment, what “support” means in this BM study should be clarified.  We don’t agree with the comments of having Alt 3 as a bench mark or something. The baseline or bench mark should be legacy non-AI approaches. All these three Alts are actually different schemes which potentially requires spec support and AI functionality. Further study of these three Alts should include aspects like performance, RS overhead, UE complexity, etc. We prefer to add a note to clarify this, like following.  ***Note: Further study of these three alternative schemes should include aspects like performance, RS overhead, UE complexity, etc.*** |
| OPPO | Support the FL proposal in principle.  In our understanding, Alt.1 and Alt.2 involves the feature of spatial domain beam prediction, whereas Alt.3 are pure temporal domain beam prediction on which BM-Case2 should focus. It is reasonable to mark Alt.3 as a baseline and allow other alternative (mixed spatial/temporal domain) as agreed as well. |
| qualcomm | Agree with inclusion of Alt.3 as the baseline and prioritizing it for study. Alt. 1 and Alt. 2 can be considered in later phases of the study item. |
| HW/HiSi | Support. However, we have our doubts on Alt 3, since for BM-Case 2, if the set A and set B are the same, there is no gain in overhead reduction. During the observation phase, the number of beams that needs to be swept is too large. Then, the whole gain with temporal BM would be lost and one could also simply do spatial domain BM more often. |
| NTT DOCOMO | Support the proposal. Our preference is to consider Alt.3 as baseline and Alt1/2 as optional, because Alt.3 can check the gain of AI/ML model based on only temporal beam prediction without the gain of spatial domain beam prediction. |
| Ericsson | Support. Except that we don’t see a need for including the notes on wide and narrow beam. We propose to remove such terms. |
| Intel | We think Alt-3 should be the baseline and the study should first focus on Alt-3. Overhead reduction with other alternatives can be studied after initial evaluations. |
| Apple | Support the proposal. |
| Samsung | Similar wording suggestion as above. Also, we tend to agree that Alt.3 can be regarded as baseline to see the performance gain of AI-based beam prediction in the case of high UE speed.  ***Proposal 2.2.2-2: For the sub use case BM-Case2, support the following alternatives for further study:***   * ***Alt.1: Set A and Set B are different (e.g. Set A consists of narrow beams and Set B consists of wide beams)*** * ***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: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.*** * ***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.*** |
| CMCC | For BM-Case2, we think Alt.3 can be the starting point. |
| Mod | According to the inputs, most companies (LGE, CATT, ZTE MTK, NEC, Lenovo, CAICT, NVIDIA, Xiaomi, Panasonic, OPPO, QC, DCM, Intel, Samsung, CMCC) support to take Alt.3 as a baseline and Alt.1/2 can be optional or discussed in a later stages. In contrast, some companies (Google, vivo, HW) don’t think Alt.3 should be the baseline.  The updated proposal is focusing on some wording changes suggested by the group and the three alternatives are still kept. Please continue to discuss it. |
| vivo | Suggest to add a note to clarify principles for future discussion.  ***Proposal 2.2.2-2: For the sub use case BM-Case2, support the following alternatives for further study:***   * ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.*** * ***~~Note2: The narrow and wide beam terminology is for SI discussion only and have no specification impact~~*** * ***Note2~~3~~: The ~~codebook constructions~~ beam pattern of Set A and Set B can be clarified by the companies.*** * ***Note3: Further study of these three alternative schemes should include aspects like performance, RS overhead, UE complexity, etc.*** |
| CAICT | Fine with vivo’s wording. |
| Lenovo | Support the latest FL proposal. |
| InterDigital | Support the latest proposal from the moderator. |
| NEC | We prefer the FL’s updated proposal. |
| Samsung | We prefer the FL’s updated proposal. We don’t think the note 3 added by vivo is necessary. We should focus on the definition of these cases first. |
| CATT | Support with vivo’s suggestion. |
| Qualcomm | Support updated Proposal 2.2.2-2 (latest one). |
| HW/HiSi | Fine with vivo’s wording.  We think that Alt3 would result into too much overhead that is not necessary. There, it is good that the aspects that vivo has brought up will be taken into consideration. |
| MediaTek | Prefer the FL’s updated proposal. |
| Ericsson | Support FL’s updated proposal. We don’t agree with vivos comment, this should be part of the evaluation KPI discussion. |
| Futurewei | Support the latest FL proposal. |
| Charter | Support new generalized wording of proposal 2.2.2-2 |

###### Proposal 2.2.2-2a (Closed)

For the updated version (“a” is missed) of Proposal 2.2.2, all companies are fine with the main contents. Vivo suggested to add Note3. Some companies support Note3 whereas some other companies support the original proposal. In moderator’s understanding, the aspects (e.g., performance) are reflected by the KPI of EVM session. Moreover, for all sub use cases and alternatives, similar aspects should be considered and we don’t need to add similar notes for each proposal. Thus, it seems not critical to add Note3. Let’s check whether some companies have strong concern on Proposal 2.2.2-2a (same as the latest version by adding “a” to “2.2.2-2”)

***Proposal 2.2.2-2a: For the sub use case BM-Case2, support the following alternatives for further study:***

* ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.***
* ***~~Note2: The narrow and wide beam terminology is for SI discussion only and have no specification impact~~***
* ***Note2~~3~~: The ~~codebook constructions~~ beam pattern of Set A and Set B can be clarified by the companies.***

|  |  |
| --- | --- |
| Company | Comments |
| NTT DOCOMO | Support the proposal. |
| CATT | Support |
| PANASONIC | OK. |
| Xiaomi | Support the proposal |
| Charter | Support the generalized wording |
| Mod | The corresponding agreement is made in online session |

### Beam pattern for Set B

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 6: For BM-Case 1, for the definition of Set B and Set A, for Alt. 1, a fixed pattern can be regarded as the starting point.* |
| Vivo[4] | *Proposal 1: Suggest to study subset selection method if fixed beam subset is used for AI input.* |
| OPPO[11] | *Proposal 2: For BM-Case1, Set B can be a subset of Set A with fixed pattern.* |
| CATT[13] | *Proposal 3: For the Alt.1 of sub use case BM-Case1, i.e., Set B is a subset of Set A, both fixed pattern and random pattern can be further studied to determine Set B out of the beams in Set A.*   * *FFS: How to select the fixed pattern in reality.* |
| NEC[14] | *Proposal 2: For Alt.1 in BM-Case1, support using the following beam patterns to determine Set B out of the beams in Set A: fixed pattern or random pattern.* |
| Spreadtrum[18] | *Proposal 2: For Alt.1 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.*   *Proposal 3: For sub use cases BM-Case2, evaluate and further study Alt3 as high priority.* |
| CAICT[20] | *Proposal 3: For spatial-domain beam prediction at UE side, Set B should be a subset of Set A. Set B is randomly chosen as baseline.* |
| Nokia[25] | *Proposal 2: Further compare the beam prediction performance/tradeoff between training and testing model with fixed Set B and training and testing model with randomized Set B.* |
| Charter[31] | *Proposal 1: For Set B being a subset of set A, the beam pattern of set B can be obtained via a dynamic pattern* |

When Set B is a subset of Set A, there are different alternatives on how to determine the beam pattern of Set B and the corresponding views are summarized as below:

|  |  |
| --- | --- |
| Beam pattern for Set B if Set B is a subset of Set A | |
| Fixed pattern | Huawei[2], OPPO[11], CATT[13], NEC[14], Spreadtrum[18], Nokia[25] |
| Random pattern | vivo[4], CATT[13], NEC[14], CAICT[20], Nokia[25] |

According to the tdocs, some companies suggest to do more study/evaluation to determine the beam pattern of Set B. Moderator feels that it is a good suggestion and we can further study this issue. Meanwhile, in EVM session, there are also many tdocs show the evaluation results for different alternatives. Duplicated discussion in the two sub agenda items should be avoided.

**Moderator recommendation**: In order to avoid the duplicated discussion, discuss this issue in EVM session (Agenda item 9.2.3.1).

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Fine to discuss in 9.2.3.1 |
| CATT | Fine with Moderator recommendation. |
| ZTE | Support to discuss it in Agenda 9.2.3.1. |
| MediaTek | Support |
| NEC | Support |
| CAICT | Support |
| Spreadtrum | Support |
| vivo | OK |
| OPPO | Support |
| Qualcomm | Support |
| Fujitsu | support |
| HW/HiSi | Ok to move to 9.2.3.1 |
| NTT DOCOMO | Fine with the proposal.  Once the beam pattern of Set B is determined and performance gain is observed in 9.2.3.1, the potential specification impacts related to them should be discussed in 9.2.3.2.  Mod: Agree |
| Samsung | Fine to discuss in 9.2.3.1 |
| CMCC | Agree |
| InterDigital | Fine |

## Input of BM-Case1 and BM-Case2

In RAN1#109e meeting, the agreements on the input of AI/ML modes for BM-Case1 and BM-Case2 were made as below:

|  |
| --- |
| 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  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. |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Observation 2: Input and output are essential parts of AI/ML model training/inference. It is convenient to carry out the use case discussion if the potential main input and output options are shared among companies while the details like format, shape may be considered as implementation dependent.*  *Proposal 2: Unless there is specification impact, the exact input and output for the AI/ML model (e.g., format, shape) should not be fixed or specified while potential input/output options may be discussed/shared for (sub) use case discussion purpose.* |
| Huawei[2] | *Proposal 7: For input to the AI/ML model, to study the spec impact, performance gain and feasibility*   * *Consider Alt1 as baseline since it is simple and can already provide considerable performance.* * *Companies may report other input to the AI/ML model according to Alt 2, 3 or 4* |
| vivo[4] | *Proposal 3: Assistance information, such as Tx/Rx beam ID or angle in connection with input RSRPs, should be used as AI input with random subset selection for both BM-case1 and case2.*  *Proposal 4: Suggest to use both Tx and Rx beam information as assistance information for further performance improvement in random subset selection.*  *Proposal 5: Study semi-random beam subset scheme with Tx/Rx beam information as AI input for both BM-case1 and BM-case2.*  *Proposal 7: Study expected information method as the input as one of the solutions for generalization to different number of Tx/Rx beams in BM-case1.*  *Proposal 8: Further study expected information method in BM-case2.*  *Proposal 9: Further study multiple expected beam information simultaneously used in AI input.*  *Proposal 14: Further study assistance information, such as beam shape pattern, 3dB beam width, etc., as model input to address performance deterioration for generalization of different beam shapes in both BM case-1 and BM case-2.* |
| ZTE[5] | *Proposal 4: Focusing the discussion on Alt.1 and Alt.4 with only L1-RSRP measurement and the corresponding beam ID being taken into account for the AI input would be a good starting point. (BM-Case1)*  *Proposal 5: Focusing the discussion on Alt.1 and Alt.2 as the starting point. The corresponding relationship between the output beam direction or angle and the TCI state needs to be further studied if Alt.3 or Alt.4 is adopted as the AI output. (BM-Case2)*  *Proposal 7: For temporal domain beam prediction, focusing the AI input and output on measured RSRP and/or beam ID would be a good starting point, in which case the standardization workload and AI model complexity would be relatively low.* |
| IDC[8] | *Proposal 6: Support ‘L1-RSRP measurement based on Set B and the corresponding DL Tx and/or Rx beam ID’ as a baseline.*  *Proposal 7: Additional information such as TRP IDs and Panels IDs should be considered.*  *Proposal 8: ‘CIR based on Set B’ can be considered as an alternative only for beam management based on FR1 information.* |
| Google[9] | *Proposal 1: For spatial domain beam prediction, support Alt3 (CIR based on set B).*  *Proposal 2: For spatial domain beam prediction, support to add CIR+L1-SINR as one alternative, where the L1-SINR can be used to reflect the interference level for the CIR measurement.*  *Proposal 5: For time-domain beam prediction, support to add CIR measurement based on set B as one alternative.*  *Proposal 6: For time-domain beam prediction, support to add CIR+L1-SINR as one alternative, where the L1-SINR can be used to reflect the interference level for the CIR measurement.* |
| OPPO[11] | *Proposal 3: For BM-Case1, whether/how the DL Tx and/or Rx beam IDs are input should be clarified.*  *Proposal 4: For the assistance information of BM-Case1, suggest to*   * *Justify the performance benefits if assistance information applied* * *Study whether assistance information would expose beamforming implementation and proprietary information at any side*   *Proposal 7: For BM-Case2, whether/how the DL Tx and/or Rx beam IDs are input should be clarified.*  *Proposal 8: For assistance information of BM-Case2, suggest to*   * *Justify the performance benefits when assistance information input to model* * *Study whether assistance information would expose beamforming implementation and proprietary information at any side* |
| BJTU[12] | *Proposal #3: Consider using wide beams and related RSRP measurements as well as extra information such as UE position and speed as input of the AI/ML model. Consider using the narrow beam RSRP prediction as the output of the AI/ML model.* |
| CATT[13] | *Proposal 5: For the sub use case BM-Case1, the following alternatives can be considered 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.*  *Proposal 10: For the sub use case BM-Case2, the following alternatives can be considered 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.* |
| NEC[14] | *Proposal 3: For BM-Case1, assistance information in input should be discussed in different deployments of AI/ML model, i.e., at gNB only, at UE only.*  *Proposal 6: For BM-Case2, assistance information in input should be discussed in different deployments of AI/ML model, i.e., at gNB only, at UE only.* |
| Lenovo[15] | *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.* |
| NVIDIA[16] | *Proposal 2: For BM-Case 1, at least support L1-RSRP measurement based on Set B of beams as AI/ML model input.*  *Proposal 3: 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 4: 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.*  *Proposal 5: Comprehensive evaluation results showing convincing performance gains is needed to nail down the essential assistance information needed for the temporal DL beam prediction.* |
| CAICT[20] | *Proposal 6: L1-RSRP and DL Tx and/or Rx beam ID could be considered as AI model input for both time domain and spatial domain beam prediction.* |
| LGE[22] | *Proposal #2: For the UE AI/ML input, Alt2 can be considered including assist information, e.g. beam grid information.* |
| Ericsson[24] | *Proposal 3 Assistance information related to “beams” should focus on information related to NW antenna/beam configuration ID or UE antenna/beam configuration ID*  *Proposal 4 Prioritize assistance information that can be obtained with low standardization effort, such as UE position information*  *Proposal 6 Investigate assistance information that capture time-dynamics without requiring any L1-RSRP measurements over a long time duration* |
| Nokia[25] | *Proposal 4: Further study the use of assistance information for ML model input to the NW side to improve DL Tx beam prediction, and the mechanism for acquiring such information through air-interface.*  *Proposal 9: RAN1 further studies the use of assistance information for ML model input to the UE side. Assistance information may include the UE’s angle relative to a panel array of the gNB and the beam boresight direction for the measured DL Tx beams to improve DL Tx beam prediction.*  *Proposal 14: For BM-Case1 with Set A/B consider 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.* |
| MTK[26] | *Proposal 3: RAN1 will discuss and agree on the alternatives for AI/ML input for Spatial Domain Beam Prediction (BM-Case1).*  *Proposal 5: RAN1 will discuss and agree on the alternatives for AI/ML input for Temporal Domain Beam Prediction (BM-Case2).*  *Proposal 6: RAN1 will study on the details and advancement of UE’s beam-related L1-RSRP report.*  *Proposal 7: Discussions and agreements are needed to prioritize and down-scope alternatives of UE assistance information.* |
| Apple[28] | *Proposal 1: clarify the Alt. 1 and Alt. 4 for use case 1 and alt. 1 and Alt. 3 for use case 2.*  *Proposal 1a: study the use of CIR for AI aided BM.*  *Observation 1: 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 2: 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.*  *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[29] | *Proposal 7: Support mechanisms to provide DL Tx beam information from NW to UE for DL beam prediction with UE side model, if it is beneficial for the beam prediction with UE side model.* |

According to the tdocs submitted to this meeting, companies’ views are quite diverging, especially for the assistance information. Thus, moderator’s tentative suggestion is to further discuss these issues and encourage the proponents to provide more details/show benefits to convince other companies.

###### Proposal 2.3 (Placeholder)

***Proposal 2.3***(TBD)

|  |  |
| --- | --- |
| Company | Comments |
| LGE | To our understanding, alt 4 can be viewed as a subset of alt 2 since beam ID can be considered as an assist info.  Mod: In the last meeting, some companies Alt.4 is the basic one and should be highlight lied explicitly. Thus, it is agreed as a separate alternative. |
| CATT | At least Alt.1 is supported by most of companies. Maybe we can have some conclusion on Alt.1 as following.  For assistance information, we can discuss it based on the evaluation results in EVM session or further discuss based on more evaluation results in the next meeting.  ***Proposal 2.3a: For the sub use case BM-Case1 and BM-Case2, the AI/ML input at least includes the following alternatives:***   * ***Alt.1: Only L1-RSRP measurement based on Set B*** * ***FFS on assistance information and Beam ID*** |
| ZTE | The current candidates for assistance information are quite diverging. According to the evaluation results provided by most of companies, focusing the AI input on the measured RSRP and/or beam ID would be a good starting point. The measured RSRP and beam ID represent the beam quality and beam indicator, respectively, where the beam ID can be implicitly indicated by the RS index or TCI state. In this way, the standardization workload and AI model complexity would be relatively low. |
| Google | We think the down-selection requires more study. But if possible, we suggest having clear sub-alternatives under Alt2. Otherwise, it is hard to study Alt2. |
| Spreadtrum | We support Alt 1. We think whether to support Alt 3 is depended on the pattern of Set B |
| vivo | We are generally okay to discuss this issue starting from the two conclusions we achieved in last meeting. Further details on what assistance information is needed can be the focus of this discussion. |
| Sony | Support IDC’s proposal on using FR1 information for prediction in FR2 |
| Fujitsu | RSRP and/or beam ID can be a starting point. For other assistance information, it can be FFS and companies need to show the gains when evaluation results are submitted. |
| HW/HiSi | We prefer Alt1. Since assistance information is highly related to the random Set B, and multiple companies have found that the assistance information does not provide gain for fixed pattern Set B, we prefer to discuss this topic in 9.2.3.1 as well. We are wondering why the assistance information is needed and if proponents could show possible gains. |
| NTT DOCOMO | In EVM session, the potential gain by assistance information should be discussed. Once the gain is observed, we should consider the potential impacts in this session. |
| Samsung | We have a question on Alt-1. If Alt-1 is adopted, does it mean that the L1-RSRP corresponding to the same set of beam indexes will be regarded as AI input (so that no explicit beam index information is needed)? Some clarification on the agreements made in previous meeting is needed. |
| Mod | The proponent(s) of Alt-1 is encouraged to make some clarification on Samsung’s question.  Based on the inputs, it seems Alt.1 and Alt.4 can be acceptable to most companies. Meanwhile, some companies suggest to clarify the assistance information and study the benefits.  The proponents of Alt.3 are encouraged to share the detailed schemes and the benefits to convince other companies. |
| Vivo | We support Alt 2 and Alt 4 for BM Case 1, and Alt 2 and Alt 3 for BM Case 2.  We think Alt 1 itself has many issues to be solved. Our understanding of Alt 1 is that it uses one single fixed set B. Then the performance can be degraded heavily if some beams in set B suffer blockage or cause large inter-cell interference to neighbor cells. Assistant information is needed to solve these issues. This aspect requires careful study to make sure the outcome of AI beam can be useful in real deployment. |
| NEC | For BM-Case1, we prefer Alt.1, Alt.2 and Alt.4. For BM-Case2, we are OK for all alternatives. Furthermore, for BM-Case1 or BM-Case2, the corresponding alternatives depends on whether model inference is performed at NW side or UE side, so we suggest adding the following note, i.e., Note5 in Proposal 2.4a proposed by Qualcomm.   * ***NoteX: All of the inputs in the above alternatives may vary based on whether the AI/ML model inference is at UE side or gNB side.*** |
| ZTE | Regarding Samsung’s question, we think Alt 1 indeed means that only L1-RSRP measurement based on Set B with fixed beam pattern is input to the AI model. The beam index information can be considered to be self-included, which can be implicitly obtained from the beam pattern. However, for a random beam pattern, the associated beam ID may have to be used as an assistance information for the AI input to improve the beam prediction accuracy. |

## Output of BM-Case1 and BM-Case2

Regarding the output of BM-Case1 and BM-Case2, there were intensive discussions and several versions of proposals were proposed. Unfortunately, no consensus was achieved in the last meeting. The final versions of the corresponding proposals were as below:

|  |
| --- |
| ***Proposal 2-4d: Regarding the sub use case BM-Case1, further 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 predicted Top-N1 DL Tx and/or Rx beams***    + ***FFS: how to select Top-N1 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.***) * ***Alt.2: Tx and/or Rx Beam ID(s) of the predicted Top-N1 DL Tx and/or Rx beams and other information***    + ***FFS: other information (e.g., probability for the beam to be the best beam, an updated set B)*** * ***Alt.3: The predicted RSRP corresponding to the ~~expected~~ Tx and/or Rx beam direction which is input to the model.*** * ***Alt.4: Tx and/or Rx Beam angle(s) and the predicted L1-RSRP (optional) of the predicted Top-N1 DL Tx and/or Rx beams*** * ***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 N1 is up to each company.***   ***Proposal 3-5c: Regarding the sub use case BM-Case2, further study the following alternatives for AI/ML output (one prediction for a future time instance) with potential down-selection:***   * ***Alt.1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the predicted Top-N2 DL Tx and/or Rx beams***    + ***FFS: how to select Top-N1 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.***) * ***Alt.2: Tx and/or Rx Beam ID(s) of the predicted Top-N2 DL Tx and/or Rx beams***    + ***FFS: other information (e.g., probability for the beam to be the best beam, the associated confidence)*** * ***Alt.3: Tx and/or Rx Beam angle(s) and/or and the predicted L1-RSRP of the predicted Top-N2 DL Tx and/or Rx beams*** * ***Alt.4: The predicted RSRP corresponding to the expected Tx and/or Rx beam direction and expected timing occasions which are input to the model.*** * ***Alt.5: Tx and/or Rx Beam ID(s) and the corresponding beam application time/dwelling time*** * ***Alt.6: Predicted Beam failure and the corresponding Tx beam ID(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*** |

The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Observation 2: Input and output are essential parts of AI/ML model training/inference. It is convenient to carry out the use case discussion if the potential main input and output options are shared among companies while the details like format, shape may be considered as implementation dependent.*  *Proposal 2: Unless there is specification impact, the exact input and output for the AI/ML model (e.g., format, shape) should not be fixed or specified while potential input/output options may be discussed/shared for (sub) use case discussion purpose.* |
| ZTE[5] | *Proposal 5: Focusing the discussion on Alt.1 and Alt.2 as the starting point. The corresponding relationship between the output beam direction or angle and the TCI state needs to be further studied if Alt.3 or Alt.4 is adopted as the AI output.*  *Proposal 7: For temporal domain beam prediction, focusing the AI input and output on measured RSRP and/or beam ID would be a good starting point, in which case the standardization workload and AI model complexity would be relatively low.* |
| Sony[6] | *Proposal 1: In output of AI/ML, indicate the evaluate criteria associated with the predicted beam ID in BM-case1 and BM-case2 for example TX beam ID for maximum dwelling time, TX/RX beam ID for maximum RSRP, etc.*  *Proposal 2: BM-case2: AI/ML output a set of Tx and/or Rx beams for a sum probability of being the best beams higher than a threshold.* |
| Google[9] | *Proposal 3: For spatial domain beam prediction, support the best beam possibility for each beam in Set A as the output.*  *Proposal 7: For time-domain beam prediction, support the best beam possibility for each beam in Set A as the output.* |
| OPPO[11] | *Proposal 5: For the output of AI/ML model for BM-Case1, 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*   *Proposal 9: For the output of AI/ML model for BM-Case2, suggest to include*   * *Tx and/or Rx Beam ID(s) for F time instances* * *The predicted L1-RSRPs of the predicted Top-K DL Tx and/or Rx beams for F time instances* |
| BJTU[12] | *Proposal #3: Consider using wide beams and related RSRP measurements as well as extra information such as UE position and speed as input of the AI/ML model. Consider using the narrow beam RSRP prediction as the output of the AI/ML model.* |
| CATT[13] | *Proposal 6: For the sub use case BM-Case1, the AI/ML outputs at least include:*   * *Alt.1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the predicted Top-N1 DL Tx and/or Rx beams;*   + *Both Top-N1 L1-RSRP and/or Top-N1 sum probability of being the best beams can be used to select Top-N1 DL Tx and/or Rx beams;*   + *Tx and Rx Beam ID(s) is indicated by using SSBRI or CRI;*   + *Values of N1 can be 1, 2, 3 or 4.* * *Alt.2: Tx and/or Rx Beam ID(s) of the predicted Top-N1 DL Tx and/or Rx beams and other information.*   + *FFS: other information (e.g., probability for the beam to be the best beam, an updated set B).*   *Proposal 11: For the sub use case BM-Case2, the AI/ML outputs at least include:*   * *Alt.1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the predicted Top-N2 DL Tx and/or Rx beams;*   + *Both Top-N2 L1-RSRP and/or Top-N2 sum probability of being the best beams can be used to select Top-N2 DL Tx and/or Rx beams;*   + *Tx and Rx Beam ID(s) is indicated by using SSBRI or CRI;*   + *Values of N2 can be 1, 2, 3 or 4.* * *Alt.2: Tx and/or Rx Beam ID(s) of the predicted Top-N2 DL Tx and/or Rx beams;*   + *FFS: other information (e.g., probability for the beam to be the best beam, the associated confidence).* * *Alt.5: Tx and/or Rx Beam ID(s) and the corresponding beam application time/dwelling time;* * *Alt.6: Predicted Beam failure and the corresponding Tx beam ID(s).* |
| NEC[14] | *Proposal 4: For Alt.1, 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 beams higher than a threshold, L1-RSRP higher than a threshold.* |
| Xiaomi[19] | *Proposal 2: For BM-Case2, the periodicity of future time instance can be same or shorter than that of history measurement instance.* |
| Ericsson[24] | *Proposal 5 No need to define the exact ML-model output for spatial beam predictions, model output should be part of the model description when presenting the simulation results*  *[Like the spatial beam prediction, there is no need to define the exact ML-model output as long as the models are evaluated with same KPI metrics.]* |
| Nokia[25] | *Proposal 1: Regarding the sub-use case BM-Case1, further study the following alternatives for AI/ML output:*   * *Option 1: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the predicted Top-N1 DL Tx and/or Rx beams*    + *For L1-RSRP prediction, the N1 selection threshold should depend on the measurements from Set B*   + *For Beam ID prediction, N1 should be a fixed value.*   + *FFS: the value for and N1.* * *Option 2: Tx and/or Rx Beam ID(s) of the predicted Top-N1 DL Tx and/or Rx beams and other information*    + *The other information can be used to derive Top-N1 DL Tx and/or Rx beams or considered as additional information*   + *FFS: other information (e.g., a QoS based metric, beam angles)*   *Proposal 17: For BM-Case2, as model output, RAN1 further discusses the detail of the prediction confidence level.*  *Proposal 18: Regarding the sub-use case BM-Case2, the AI/ML output should consider:*   * *Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the predicted Top-N2 DL Tx and/or Rx beams*    + *For L1-RSRP prediction, the N1 selection thresholdshould depend on the measurements from Set B*   + *FFS: the value of N2.*   *Proposal 20: For BM-Case2 model inference in UE side, NW may configure UE to report the related prediction quantity (i.e. confidential level, RSRP error, observation window length), as well as the predicted beams for one or more future instants.* |

Based on the submitted tdocs, Alt.1 and Alt.2 are supported by most companies. There is also some company supporting not to define any output. Meanwhile, some company(es) suggest to reduces the number of alternatives. Taking the afore-mentioned information into account, moderator suggests to take the following proposal as a starting point, which is modified from Proposal 2-4d and Proposal 3-5c of RAN1#109e meeting:

* Proposal 2-4d of RAN1#109e meeting is modified to Proposal 2.4
  + Alt.3 is merged to Alt.1
* Proposal 3-5c of RAN1#109e meeting is modified to Proposal 2.4
  + Alt.4 is merged to Alt.1
  + Alt.5 is merged to Alt.2
  + Alt.6 is merged to Alt.2

Proposal 2.4

***Proposal 2.4: Regarding the sub use case BM-Case1 and BM-Case2, support to 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 predicted Top-N DL Tx and/or Rx beams*** 
  + ***FFS: how to select Top-N1 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)***)
* ***Alt.2: Tx and/or Rx Beam ID(s) of the predicted Top-N 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)***
* ***Alt.3: Tx and/or Rx Beam angle(s) and the predicted L1-RSRP (optional) of the predicted Top-N DL Tx and/or Rx beams***
* ***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.***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | We are not sure whether/how output of AI/ML model impacts specification, which may still be within implementation domain.  Mod: It is for discussion purpose as explicitly highlighted by Note3. If there is no clear scheme for the discussion, there will be much confusion for the group. Whether/how the inputs impact the spec, it is up to the further discussion. |
| CATT | Support this proposal.  The output will impact the UE reporting and UE procedure after the AI/ML prediction, if AI/ML is deployed in UE side. If AI/ML is deployed in NW side, the output will impact the configuration. Thus, it’s necessary to discuss the AI/ML model outputs. |
| ZTE | Support the proposal. We agree with CATT that the AI output will impact the reporting and resource configuration, and thus needs to be discussed. For instance, if beam ID is predicted by the UE-side model in Alt.2, only beam ID needs to be reported without the associated RSRP, which is different with the current spec. |
| MediaTek | Support this proposal. Down-scoping of “other information” could be a viable way to study forward. |
| NEC | Support |
| Lenovo | We understand that the goal of AI for BM is to study the possible of adopt AI based beam prediction to replace of Rel-17 beam measurement procedure for overhead/latency reduction. So, Rel-17 beam reporting should be the baseline for AI/ML output at least for UE-centric AI inference. For NW-centric AI inference, it seems AI/ML output does not have spec impact.  Mod: we can study it and make some conclusion (if possible) later |
| CAICT | Support |
| NVIDIA | Support to further study Alt. 1 and Alt. 2. |
| Futurewei | We are ok with the proposal. |
| Google | Suggest adding the following Alt4. In addition, shall we change “support to study” into “study” in the main-bullet?   * ***Alt.4: Tx and/or Rx angle(s) of the predicted Top-N 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)***   Mod: Would you like to elaborate a bit more on how to determine the angle for a practical system, especially for a UE ? |
| Xiaomi | Support this proposal and prefer Alt 1. Clarification on other information is needed for Alt 2. |
| Spreadtrum | Support |
| vivo | We are not clear about what the intention or the point is to discuss AI output. Is it for EVM purpose or spec impact purpose?  We think this agenda item should focus on spec impact purpose. Then in this regard, we tend to agree with Ericsson that we shouldn’t define output from spec impact perspective, as they are based on implementations in most of the cases. Output can just be reported by companies when presenting their results. Hence our suggestion is first to discuss what the relationship between AI output and spec impact is, and then discuss potential categorizations.  Mod: If we are to discuss the spec impact, we need to have some assumed scheme in mind. If we don’t have some any clear scheme, it would be difficult to study the spec impact. The intention of this proposal to clarify/list the typical schemes in the mind of companies to facilitate the further discussion, e.g., encourage companies to provide evaluation results for focused schemes, study spec impact with focus. Moreover, Note3 clarified it is for discussion purpose |
| Sony | Support both Alt 1 and 2. |
| OPPO | Support.  In our view, the output of AI/ML model has standard impact over beam measurement, reporting or beam indication in NR. |
| qualcomm | Support the proposal in principle and agree that AI/ML model outputs should be discussed to identify the specification impact. |
| Fujitsu | Support |
| HW/HiSi | Support |
| NTT DOCOMO | We are fine with capturing all the potential outputs even though the output is not directly captured in the specification. If the expected output format is limited as proposal, the discussion can be facilitated because companies assumes the same AI/ML output formats in common. |
| Ericsson | No need to define the exact ML-model output for beam predictions, model output should be part of the model description when presenting the simulation results.  Agree with Vivo that first we need to identify the spec impact of the ML-model output.  Mod: please see the reply to vivo |
| Intel | Tend to agree with companies that this should be more about spec impact.  Mod: please see the reply to vivo |
| Samsung | In our understanding, there are too many combinations in each alternative which is not preferable for the evaluation/justification in 9.2.3.1. We suggest to further narrow down those alternatives to popular sub-alternatives so that we can focus on them first for evaluation and specification impacts.  Moreover, FFS under each sub bullet should be deleted. They are next level of discussion and shall not be the focus.  ***Proposal 2.4: Regarding the sub use case BM-Case1 and BM-Case2, support to study the following alternatives for AI/ML output:***   * ***Alt.1: Top 1 or N Tx beam ID(s)*** * ***Alt 2: Top 1 or N Rx beam ID(s)*** * ***Alt 3: Top 1 or N Tx beam and Rx Beam ID(s) or Tx-Rx beam pair ID(s)*** * ***FFS on other information other than beam ID(s) at least including the predicted L1-RSRP***   ***FFS on number of N***  Mod: let’s check other companies’ view on the down-selection. It would be better if the group can achieve consensus on some down-selection. |
| CMCC | Support. If the AI/ML inference is deployed at UE side, the necessary information for reporting should be considered. |
| Mod | Please continue to discuss it, including the down-selection proposed by SS. |
| vivo | It is still confusing to us why we have to define exact AI output in spec impact discussion. For example, all these alternatives talk about top N beams as the output. Why does the AI output have to be top N beams? Can’t the implementation just be to predict all the beams or all the beams required by gNB or UE as AI output, and then select the best ones for future utilization (e.g., TCI indication or beam reporting) ? Considering this aspect, isn’t it more important to discuss what is configured/indicated by gNB and what is reported by UE? |
| CAICT | Support |
| Fujitsu | support |
| InterDigital | In our view, AI/ML does not need to select Top 1 or N Tx/Rx beam. Actually, AI/ML model can predict L1-RSRP value for each Tx-Rx beam pair and gNB or UE can select a best beam by its implementation. Having said that, we propose the following alternative.   * ***Alt.4: Tx and/or Rx Beam ID(s) and/or the predicted L1-RSRP of the predicted DL Tx and/or Rx beams*** |
| NEC | Support |
| CATT | Support this proposal. For the potential down-selection, we prefer at least support Alt.1 and Alt.2. |
| Qualcomm | All of the above alternatives depend on whether the AI/ML is UE-side or gNB-side. That is why we suggest adding the following note:   * ***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.*** |
| Apple | We have “alt. 3 f or BM case 1, we suggest add back “CIR” as another alternative. |
| Ericsson | We first note that the agreement only considers “AI/ML output”, and not AI/ML model output”,  We share Vivos view, one source of confusion is the unclarity of what could constitute a “model”. Assume the model predicts all beams, then we might have some non-ML logic (postprocessing) that selects the “best beams” based on such predictions. In the end there will come out a top-N beam IDs, but it is not clear if this is from the “model” output, or from the non-ML logic processing the output.  The problem of what constitutes a model output is even more apparent in the FFS note “*FFS: how to select Top-N1 DL Tx and/or Rx beams…”* . The term “select” assumes there is something non-ML that selects beams. How would you define a loss function for such scenario otherwise?  We still don’t think it is needed to define the exact output of the model, but since our understanding is that the agreement considers “AI/ML output”, and companies assume there is a non-ML logic post-processing the actual ML-model output to derive e.g. top-N beams. We can accept the proposal but add the note below.  ***Note 6: The Top-N1 beam IDs might have been derived via post-processing of the ML-model output”*** |
| Charter | Support to further study Alt 1 and Alt 2 |

Proposal 2.4a

IDC and QC inputs some proposal alternative/note for this proposal. In moderator’s understanding, Alt.4 is covered by Atl.1 since N can be the number of all beams or beam pairs. Moreover, Note 6 can also address the concern of IDC in some senses. Thus, in the updated version, Alt.4 is not included. Please feel free to correct me if I missed something.

***Proposal 2.4a: Regarding the sub use case BM-Case1 and BM-Case2, support to 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 predicted Top-N DL Tx and/or Rx beams*** 
  + ***FFS: how to select Top-N~~1~~ 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)***)
* ***Alt.2: Tx and/or Rx Beam ID(s) of the predicted Top-N 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)***
* ***Alt.3: Tx and/or Rx Beam angle(s) and the predicted L1-RSRP (optional) of the predicted Top-N DL Tx and/or Rx beams***
* ***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”***

|  |  |
| --- | --- |
| Company | Comments |
| Futurewei | Support the updated proposal. |
| NTT DOCOMO | Support the proposal. |
| NEC | Support the updated proposal. |
| vivo | Based on FL’s reply in the previous round, Note 4 and Note 6, the so-called “top N” beams do not exactly mean to have top N beams generated from the AI/ML model. It can be all the RSRPs, or some specific RSRPs based on UE or gNB’s need and AI implementation. Is that correct?  If so, why can’t we simply remove “Top N” in the alternatives, as it does not give much information? |
| CATT | Support the FL’s updated |
| Xiaomi | Support the proposal with note 6. |

###### Proposal 2.4b (H)

Proposal 2.4b is updated from Proposal 2.4a by removing the “Top-N” as suggested by vivo. Since “Top-N” is removed, Note 6 becomes unnecessary. Thus, Note 6 is also removed. Let’s check whether it is acceptable to all companies.

***Proposal 2.4b: Regarding the sub use case BM-Case1 and BM-Case2, support to 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 ~~Top-N~~ DL Tx and/or Rx beams*** 
  + ***FFS: how to select the N ~~Top-N1~~ 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)***)
* ***Alt.2: Tx and/or Rx Beam ID(s) of the N predicted ~~Top-N~~ 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)***
* ***Alt.3: Tx and/or Rx Beam angle(s) and the predicted L1-RSRP (optional) of the N predicted ~~Top-N~~ DL Tx and/or Rx beams***
* ***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”~~***

|  |  |
| --- | --- |
| Company | Comments |
| Futurewei | We are ok with the updated proposal 2.4b. |
| ZTE | Support the updated proposal. |
|  |  |
|  |  |
|  |  |
|  |  |

## Use cases

In RAN1#109e meeting, sub use cases and categories were captured in FL summary [33] as below:

|  |  |
| --- | --- |
| Category | Sub use case |
| Cat1:  Spatial-domain DL beam prediction | **BM-Case1:** Spatial-domain DL beam prediction for Set A of beams based on measurement results of Set B of beams |
| **BM-Case3:** Beam prediction for higher frequency band (e.g., a band in FR2) based on measurement results of lower frequency band(s) (e.g., a band in FR1) |
| **BM-Case4:** Beam prediction based on UE positioning/trajectory |
| **BM-Case6:** Spatial-domain UL beam prediction for Set A of beams based on measurement results of Set B of beams |
| **BM-Case9:** Joint DL/UL beam pair link prediction |
| Cat2:  Time-domain DL beam prediction | **BM-Case2:** Temporal DL beam prediction for Set A of beams based on the historic measurement results of Set B of beams |
| Cat3: Others | **BM-Case7:** beam measurement feedback compression |
| **BM-Case8:** The beam-specific parameter optimization |

There are some discussions on these sub use cases in the tdocs. The related proposals/ observations are copied as below:

|  |  |
| --- | --- |
| Huawei[2] | *Proposal 1: RAN1 should focus on the evaluation of BM-Case 1 and BM-Case 2. Other use cases should not be included.* |
| TCL[3] | *Proposal 1: The UE position information is not necessary for predictive beam switching.*  *Proposal 2: The predictive beam switching shall be discussed in sub use cases of inter-cell beam switching and intra-cell beam switching for latency reduction.*  *Proposal 4: The beam failure detection performance can be enhanced by an AI/ML model based on historical beam measurements.*  *Proposal 5: The new candidate beam qnew can be determined by an ML model when beam failure occurs.* |
| ZTE[5] | *Proposal 1: Since the time unit of this study item is limited, we suggest to focus on the sub use cases BM-Case1 and BM-Case2 in the first phase, and whether to discuss other sub use cases depends on the progress of the first phase.* |
| Sony[6] | *Proposal 3: Support BM-case3: Beam prediction for higher frequency band (e.g., a band in FR2) based on measurement results of lower frequency band(s) (e.g., a band in FR1).* |
| OPPO[11] | *Proposal 1: Study BM-Case1 and BM-Case2 as representative sub use case with high priority.* |
| BJTU[12] | *Proposal #1: Consider high-speed railway as one of the scenarios for AI/ML based beam management. Study the implementation and design of AI/ML based beam management scheme in various railroad track scenarios.*  *Proposal #2: Support RAN1 to study the AI/ML based image super-resolution scheme for spatial-temporal beam prediction in high-speed railway scenarios as a use case for beam management enhancement, which can significantly reduce the overhead of beam sweeping.* |
| CATT[13] | *Proposal 12: For AI/ML-based beam management, the following sub use cases are deprioritized:*   * *BM-Case3: Beam prediction for higher frequency band (e.g., a band in FR2) based on measurement results of lower frequency band(s) (e.g., a band in FR1);* * *BM-Case6: Spatial-domain UL beam prediction for Set A of beams based on measurement results of Set B of beams;* * *BM-Case8: Parameter optimization to improve performance of multi-beam system;* * *BM-Case9: Joint DL/UL beam pair link prediction.*   *Proposal 13: For AI/ML-based beam management, BM-Case4, i.e., beam prediction based on UE positioning/trajectory, can be studied together with BM-Case1 and BM-Case2.*  *Proposal 14: For AI/ML-based beam management, BM-Case7, i.e., beam measurement feedback compression, can be studied similarly with the use case of CSI feedback compression.* |
| Lenovo[15] | *Proposal 3: Beam prediction at gNB/TRP side with model management-related collaboration between gNB and UE can be taken as a sub-use case for beam management in predictable trajectory scenario.* |
| NVIDIA[16] | *Proposal 1: Beam prediction in spatial domain and beam prediction in time domain should be the focal point for evaluating AI/ML based algorithms for beam management* |
| Intel[17] | *Proposal 1: BM-Case 1 can be further sub-divided into two sub-cases where Set B is either a subset of Set A or not.*  *Proposal 2: BM-Case6 should be supported for UE Tx/Rx beam prediction*  *Proposal 3: BM-Case9 should be supported since it can provide large latency and measurement gains for joint P2/P3 procedure* |
| Xiaomi[19] | *Proposal 1: For AI/ML-based beam management, only support BM-Case1 and BM-Case2.* |
| CAICT[20] | *Proposal 1: Sub use case descriptions of AI/ML-based BM could be further discussed combining with collaboration level.*  *Proposal 2: AI/ML-based time domain and spatial domain BM should be studied separately.* |
| Samsung[12] | ***Case-1b***  *This case is similar to BM-Case1 but is for UL beam prediction.*  ***Case-2b***  *This is another case for DL beam prediction. (for Rx beam prediction)* |
| LGE[22] | *Proposal #7: BM sub use cases other than BM-Case1 and BM-Case2 are deprioritized during this SI.* |
| Ericsson[24] | *[The amount of work needed for the two agreed use cases are enough for the initial stages of the study item. Any other potential use case should be down prioritized.]* |
| MTK[26] | *Proposal 8: RAN1 will discuss on prediction of top beams for a frequency band in FR2 based on the measurement results of FR1.* |
| Apple[28] | *Proposal 5: Study FR2 spatial domain beam prediction with FR1 measurements as well as CSI enhancement in FR1 to facilitate the beam prediction in FR2*  *Proposal 6: Study beam dwelling time prediction based on past measurement results as well as UE power saving schemes for beam measurement with regard to predicted beam dwelling time.* |
| DCM[29] | *Proposal 1: Prioritize the discussion of spatial-domain DL beam prediction and temporal DL beam prediction from other sub use case.* |
| KT[32] | *Proposal 1: Study sub use case of beam prediction in spatial domain with high priority.* |

Companies’ view on the other sub use cases are summarized as below

|  |  |
| --- | --- |
|  | Supporting companies |
| BM-Case3 | Sony[6], Fujitsu[7], IDC[8], MTK[26], Apple[28], |
| BM-Case4 | CATT[13], Sony[6], Lenovo[15] |
| BM-Case6 | Intel[17], Samsung[12] |
| BM-Case7 | CATT[13] |
| BM-Case8 | Charter[31] |
| BM-Case9 | Intel[17] |
| Deprioritize all other sub use cases | Huawei[2], ZTE[5], Sony[6], NVIDIA[16], Xiaomi[19], LGE[22], Ericsson[24], DCM[29], KT[32] |
| Deprioritize BM-Case3/6/8/9 | CATT[13] |

From the above table, we can see that the views on whether to support other sub use cases or not are quite diverging. The proponent(s) of other use cases is encouraged to discuss with other companies and convince them. Meanwhile, let’s make a try and check companies’ view whether some other sub use cases can be accepted in addition to BM-Case1 and BM-Case2. It seems that BM-Case3 gets more supporting companies than other sub use cases. Thus, the following proposal is suggested for further discussion.

Proposal 2.5

***Proposal 2.5: In addition to the sub use case BM-Case1 and BM-Case2, support the following sub use case(s):***

* ***BM-Case3: Beam prediction for higher frequency band (e.g., a band in FR2, or a band in FR2-2) based on measurement results of lower frequency band(s) (e.g., a band in FR1, or a band in FR2-1)***

***Conclusion 2.5:***

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

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Not support to add more sub-use-cases. |
| CATT | Not support to add BM-Case3. We can accept only study the sub use case BM-Case1 and BM-Case2 in this release. |
| ZTE | For BM-Case3, it has strong dependency on the channel information across different frequency ranges. Whether it’s reliable enough should be further studied since different frequency ranges generally have quite different channel characteristics. Besides, since the time unit of this study item is limited, we suggest to focus on the sub use cases BM-Case1 and BM-Case2 in the first phase, and whether to discuss other sub use cases depends on the progress of the first phase. |
| MediaTek | Support to include study for BM-Case 3, as it will help to predict FR2 beams using less complicated FR1 beams. |
| NEC | Not support to add BM-Case3. |
| CAICT | Not support. |
| NVIDIA | Not support. |
| FUTUREWEI | We suggest deferring other sub use(s) till later. |
| Google | Support. We have different view with ZTE, there is no dependency on channel information across frequency ranges. BM-case 3 can even work in non-co-located FR1/FR2 cells. |
| Xiaomi | Only support BM-case 1 and BM-case 2 |
| Spreadtrum | Not support. |
| vivo | We don’t support BM Case 3. We think more discussion is needed to understand the feasibility of performing such cross-band prediction considering practical impairments for hardware implementation between FR1 and FR2. |
| Sony | Agree to support Case 3. |
| qualcomm | Not support to add further use cases at this point, given the limited timeline within the scope of Rel-18. |
| Fujitsu | Not support for more sub use cases. |
| HW/HiSi | Not support |
| NTT DOCOMO | We would rather deprioritize BM-Case3 due to the workload of RAN1. |
| Ericsson | Agree to study use case 4 |
| Intel | We think joint UE-BS beam pair link prediction is an important use-case and may be treated with Case 1. This type of joint P2/P3 optimization has the potential to offer large overhead reduction benefits. |
| Apple | Support the proposal |
| Samsung | Considering the evaluation workload and spec impact, not prefer to add BM-Case3. |
| CMCC | Not support. |
| Mod | According to the inputs, although there are a limited number of companies supporting BM-Case3 or other sub use cases, 16 companies suggest to deprioritized or don’t support any additional sub use case, e.g., LGE, CATT, ZTE, NEC, CAICT, NVIDIA, FUTUREWEI, Xiaomi, Spread, vivo, QC, Fujitsu, HW, DCM, SS, CMCC  Considering the above information, a conclusion capturing the current status seems the only way to move forward. Please see the conclusion 2-5. |
| InterDigital | We support proposal 2.5 and object conclusion 2.5. |
| Qualcomm | Support Conclusion 2.5. |
| Panasonic | Support the conclusion 2.5. |
| Fujitsu | support |

###### Conclusion 2.5a (H)

According to the inputs, there are 6 companies supporting BM-Case3 or other sub use cases, whereas 16 companies suggest to deprioritized or don’t support any additional sub use case, e.g., LGE, CATT, ZTE, NEC, CAICT, NVIDIA, FUTUREWEI, Xiaomi, Spread, vivo, QC, Fujitsu, HW, DCM, SS, CMCC.

On the other side, IDC object conclusion 2.5. Meanwhile, some proponents of other sub use cases want to have more time to convince other companies.

Considering there should be enough time for EVM assumption discussion and doing evaluations, Oct meeting seems the last meeting that we can decide to support more sub use cases. Thus, moderator suggests the following conclusion. Hope it is acceptable to companies.

***Conclusion 2.5a:***

***For AI/ML based beam management, further discuss and make decision on whether or not to support any other sub use case in addition to BM-Case1 and BM-Case2 in Oct. meeting (RAN1#110b-e).***

|  |  |
| --- | --- |
| Company | Comments |
| Futurewei | We are ok with the proposal. |
| ZTE | Support |

## Spec impact

### General views

There are many contributions discussing spec impacts of AI-based beam management. The proposals/ observations related to the general principles are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 5: Regarding AI/ML-based beam management, study the standards impact, including AI/ML related UE capability reporting, related to AI/ML model selection/configuration (like activation/deactivation) in case multiple trained AI/ML models are deployed.* |
| Huawei[2] | *Proposal 8: Study potential specification impact for AI/ML-based beam prediction considering the following aspects:*   * *AI/ML model training procedure* * *Enhancement for RSRP report and beam ID report for inference* |
| vivo[4] | *Proposal 1: For both case 1 and case 2 of beam management, both collaboration level level-y-a, and collaboration level-z can be considered.*  *Proposal 20: Study specification impact on assistance information based on representative sub use cases with minimum exposures of implementation details.*  *Proposal 21: Study specification impact on beam report enhancement, especially for temporal domain beam prediction.* |
| ZTE[5] | *Proposal 8: For NW-side beam prediction AI/ML models, enhanced beam reporting mechanisms such as further screening, compression, and reporting of the beam measurement results need to be studied so as to balance the beam prediction performance and beam reporting overhead.*  *Proposal 9: For UE-side beam prediction AI/ML models, signaling methods need to be investigated to enable recommendation, configuration, and indication of the beam set for measurement.*  *Proposal 10: Study enhanced resource configuration and beam indication if more flexible triggering or activating approaches are utilized.*  *Proposal 11: Enhanced resource configuration and reporting mechanisms need to be investigated to facilitate the exchange of assistance information, which can be either implicit or explicit.* |
| Sony[6] | *Proposal 4: Propagation environment based AI/ML model selections can be considered at gNB.*  *Proposal 5: Support gNB signaling to UE in order to activate different AI/ML models at UE for beam prediction.* |
| IDC[8] | *Proposal 9: Study benefits of simple specification extension of UE reporting.*  *Proposal 10: Study benefits of specification enhancements such as UE reporting with associated time domain information.*  *Proposal 11: Study benefits of specification enhancements on association between beams with different beam widths.*  *Proposal 12: Study benefits of specification enhancements on acquiring UE Rx beam information.* |
| Google[9] | *Proposal 10: 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 11: The study of AI/ML based BM should consider both FR1 and FR2.* |
| BJTU[12] | *Proposal #4: Study potential specification impact for AI/ML-based HSR beam management, considering the following aspects:*   * + *Collaboration procedure between UE and gNB.*   + *AI/ML model deployment, training and inference procedure.* |
| CATT[13] | *Proposal 15: The following spec impact of AI/ML based beam management can be considered:*   * *Signaling/procedure of AI model training/updating/fallback;* * *Interface of AI model, i.e. relationship between measured RS and reported information;* * *New procedure for RS measurement and reporting;* * *Signaling/procedure design on exchanging AI-related/non-AI-related assistance information.* |
| NEC[14] | *Proposal 7: Study the mechanism of model update, e.g., fine-tuning.*  *Proposal 8: Study the mechanism of online data processing.*  *Proposal 9: Study the mechanism of model selection.*  *Proposal 10: Study the mechanism of reporting more beams, e.g., larger than 4.*  *Proposal 11: Study the mechanism of reducing the overhead of beam measurement and reporting in model inference.*  *Proposal 12: Study the mechanism of reducing the overhead of beam reporting in model training, model update, model testing or model monitoring.*  *Proposal 14: Study the mechanism of discontinuous reporting in periodic or semi-persistent beam reporting.* |
| Lenovo[15] | *Proposal 4: Study UE/NW capability related signaling corresponding to AI-based beam management under different network-UE collaboration levels.*  *Proposal 5: Rel-17 CSI reporting framework can be reused for NW-centric beam prediction by increasing the number of beams in a beam report.*  *Proposal 6: Rel-17 CSI reporting framework can be reused for UE-centric 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.*  *Proposal 7: How to obtain the assistant information for AI/ML input needs further study.* |
| NVIDIA[16] | *Proposal 6: 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.*  *Proposal 7: 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.*  *Proposal 8: 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.*  *Proposal 9: 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 10: 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.*  *Proposal 11: 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.* |
| Intel[17] | *Observation 2: One possible area of specification impact for AI/ML model integration may be for triggering of beam measurement reports and reference signal transmissions, as well new L1 reporting formats.* |
| Spreadtrum[18] | *Observation 2: For beam measurement and reporting, current CSI framework can be considered as starting point.*   * *If AI/ML inference is at NW side, no specification impact is identified* * *If AI/ML inference is at UE side, enhanced beam reporting needs further study*   *Observation 3: 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.*  *Observation 4: 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 the best Rx beam needs further study* * *If AI/ML inference is at UE side, no specification impact is identified* |
| Xiaomi[19] | *Proposal 6: To indicate Rx beam information to UE for obtaining L1-RSRP input to AI/ML model.*  *Proposal 7: To discuss whether a common AI model or separate AI models will be trained for UE with different number of Rx beam.*  *Proposal 8: Increase the maximum number of beams in beam report for each time instance.*  *Proposal 9: Consider enhancement on beam measurement report to contain more than one time instance.* |
| CMCC[23] | *Proposal 1: The same sort method of beam pairs is pre-defined so that gNB and UE have the same understanding of index of beam pairs.* |
| Ericsson[24] | *Proposal 7 New or enhanced mechanism(s) including CSI-report-based, SRS-based and RRC-message-based frameworks to facilitate NW data collection for beam management use cases should be studied*  *Proposal 8 Study data collection requirements and new or enhanced mechanism(s) to facilitate collecting data for NW-sided model inference for DL spatial/temporal beam prediction use cases.*  *Proposal 9 Study enhancements of CSI measurement and reporting configurations to support UE-sided DL spatial/temporal beam predictions.* |
| Nokia[25] | *Proposal 5: For UE side DL Tx beam prediction with collaboration level-y and level-z, RAN1 shall investigate further details by considering steps associated with the life cycle management of the model.*  *Proposal 6: For UE side DL Tx beam prediction, further study the necessary info required from the NW to indicate Set A and Set B relationship.*  *Proposal 7: For UE side DL Tx beam prediction, further study the RS resource set configuration for UE side DL Tx beam prediction*  *Proposal 8: For UE side DL Tx beam prediction, further study group-based beam reporting for mTRP simultaneous reception based on Set B measurements, where the UE may report beam pairs from Set A.* |
| Apple[28] | *Proposal 7: Since AI based beam prediction may not be able to provide 100% beam prediction accuracy, it is necessary to study hybrid AI based and non-AI based beam management.*  *Proposal 8: Study how to management multiple AI processing simultaneously.* |
| DCM[29] | *Proposal 2: Study the potential specification impacts of beam measurement reporting to facilitate or improve the beam prediction at NW side model.*  *Proposal 3: In DL beam prediction with NW-side model, some mechanisms to report Rx beam ID used for beam measurement can be considered as potential specification impacts.* |
| Panasonic[30] | *Observation 3: Unless Set A is the same as Set B, for AI/ML inference at UE side, the spatial relation among beams between Set A and Set B needs to be known to the UE, e.g. by specifying some rule or some signaling.*  *Proposal 1: Study how to enable the knowledge of spatial relation among beams between Set A and Set B to the UE.*  *Proposal 2: At least for the purpose of AI/ML inference at NW side, enhancement on L1-RSRP measurement configuration and reporting configuration should be considered, e.g.*   * *increasing the maximum number of reported beams* * *obtaining assistance information such as UE location, or UE Rx beam* |
| Charter[31] | *Proposal 1: Consider the option to enhance beam management with a dynamic vector–quantized codebook based on SVD and ML; it can be used and/or exchanged with the UE using e.g. a digital feedback channel between gNB and UE(s).* |
| KT[32] | *Proposal 2: Study the specification impact for both cases where the beam prediction and training functionality resides in the same or different node sides.*  *Proposal 3: Study how to signal Set B related information.* |

There are lots of high-level and detailed proposals proposed by tdocs. To roughly categorize the proposals, most of them belong to one of the following aspects:

* AI/ML Model Training
* AI/ML model inference
* AI/ML model life cycle management (LCM)
* UE capability

Thus, moderator suggest the following proposal as a starting point, which focus on the high-level aspects of potential spec impacts. Other proposals focusing on more details will be discussed in subsequent sections.

Proposal 2.6.1

***Proposal 2.6.1: For the sub use case BM-Case1 and BM-Case2, support to investigate specification impacts from the following aspects***

* ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)*** 
  + ***Note1: Online training and/or offline training is a separate discussion***
* ***New or enhanced mechanism(s) to facilitate AI/ML inference***
* ***New or enhanced mechanism(s) to facilitate AI model life cycle management***
* ***AI-related UE capability and reporting***
* ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting***
* ***Note3: Other aspect(s) is not precluded***

***Proposal 2.6.1a: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***

* ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)*** 
  + ***Note1: Online training and/or offline training is a separate discussion***
* ***New or enhanced mechanism(s) to facilitate AI/ML inference***
* ***New or enhanced mechanism(s) to facilitate AI model life cycle management***
* ***AI-related UE capability and reporting***
* ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting***
* ***Note3: Other aspect(s) is not precluded***

***Proposal 2.6.1b: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***

* ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)*** 
  + ***Note1: Online training and/or offline training is a separate discussion***
* ***New or enhanced mechanism(s) to facilitate AI/ML inference***
* ***New or enhanced mechanism(s) to facilitate AI/ML model activation/deactivation/selection/switching and fall-back operation***
* ***New or enhanced mechanism(s) to facilitate performance monitoring***
* ***AI-related UE capability and reporting***
* ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting***
* ***Note3: Other aspect(s) is not precluded***
* ***Note4: the above study should consider the associated collaboration levels***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Ok for study |
| CATT | Support |
| ZTE | Support |
| MediaTek | support |
| NEC | Support |
| Lenovo | Support |
| CAICT | Support |
| NVIDIA | Support |
| Futurewei | We are ok with the proposal. |
| Google | Just to clarify, for offline training, does it require spec impact on data collection? It is hard to use online training for BM, since even in real network, UE cannot always identify the best NW beam  Mod: It needs further discussion. This proposal does not touch this issue. |
| Xiaomi | Support. And the AI-capability can be discussed first in framework agenda. |
| Spreadtrum | Support |
| PANASONIC | Support |
| vivo | OK |
| Sony | Not sure the relationship of the life cycle management in bullet with 2.6.4   * ***New or enhanced mechanism(s) to facilitate AI model life cycle management***   2.6.4 is the detail discussion of this bullet?  Mod: Yes |
| OPPO | Support |
| qualcomm | Support |
| Fujitsu | support |
| HW/HiSi | OK |
| NTT DOCOMO | Support the proposal. |
| Ericsson | Support |
| Intel | OK in principle. Should further discuss data collection details and how it may impact spec for Offline training  Mod: Yes, we need further discussion. Section 2.6.2 is used for the detailed discussion of data collection |
| Samsung | We understand the intention of the proposal. However, for some bullet (e.g., data collection), we are not sure whether there is any specification impact. To be safe, we suggest the following modification to the proposal.  ***Proposal 2.6.1: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or the corresponding specification impacts from the following aspects.***  Mod: updated |
| CMCC | Support |
| Mod | According to the latest agreement achieved in AI 9.2.1 yesterday, the life cycle management includes all phase/aspect of AI/ML operations. In order to keep the consistency among all sub agenda item, proposal 2.6.1 and2.6.4-1 are merge as Proposal 2.6.1b |
| CAICT | Fine with the update |
| Fujitsu | Support Proposal 2.6.1b |
| Lenovo | Fine with Proposal 2.6.1b. |
| Xiaomi | Support the proposal 2.6.1b. |
| InterDigital | Fine |
| NEC | OK for Proposal 2.6.1b. Furthermore, we suggest the following modification for more clearly:  ***New or enhanced mechanism(s) to facilitate AI/ML model performance monitoring*** |
| CATT | Support Proposal 2.6.1b |
| Qualcomm | Support Proposal 2.6.1b. |
| HW/HiSi | Support the proposal 2.6.1b  ***Proposal 2.6.1b: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***   * ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)***    + ***Note1: Online training and/or offline training is a separate discussion*** * ***New or enhanced mechanism(s) to facilitate AI/ML inference*** * ***New or enhanced mechanism(s) to facilitate AI/ML model activation/deactivation/selection/switching and fall-back operation*** |
| MediaTek | Support Proposal 2.6.1b. |
| Ericsson | Support |
| Spreaftrum | Support Proposal 2.6.1b |
| Futurewei | Support Proposal 2.6.1b |
|  |  |

###### Proposal 2.6.1c (Closed)

Proposal 2.6.1b seem acceptable to all companies. One editorial change suggested by NEC and also keep the terminology consistency: performance monitoring -> AI/ML model monitoring

***Proposal 2.6.1c: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***

* ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)*** 
  + ***Note1: Online training and/or offline training is a separate discussion***
* ***New or enhanced mechanism(s) to facilitate AI/ML inference***
* ***New or enhanced mechanism(s) to facilitate AI/ML model activation/deactivation/selection/switching and fall-back operation***
* ***New or enhanced mechanism(s) to facilitate AI/ML model ~~performance~~ monitoring***
* ***AI-related UE capability and reporting***
* ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting***
* ***Note3: Other aspect(s) is not precluded***
* ***Note4: the above study should consider the associated collaboration levels***

|  |  |
| --- | --- |
| Company | Comments |
| NTT DOCOMO | Generally fine with the proposal. But there is a typo “monitoring monitoring”. The model monitoring should be written as it is captured in the current working list |
| NEC | The terminology of “Model monitoring” has been agreed in 9.2.1. So, we suggest the following:   * ***New or enhanced mechanism(s) to facilitate AI/ML ~~monitoring~~ ~~performance~~ model monitoring***   ***Mod: updated*** |
| vivo | We don’t agree to change LCM to the third sub-bullet. What LCM contains should be discussed in 9.2.1. Whatever they agree for components in LCM, we can further study the details for BM. We could be ok with the following.  ***Proposal 2.6.1c: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***   * ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)***    + ***Note1: Online training and/or offline training is a separate discussion*** * ***New or enhanced mechanism(s) to facilitate AI/ML inference*** * ***New or enhanced mechanism(s) to facilitate AI/ML model life cycle management ~~activation/deactivation/selection/switching and fall-back operation~~*** * ***New or enhanced mechanism(s) to facilitate AI/ML model ~~performance~~ monitoring*** * ***AI-related UE capability and reporting*** * ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting*** * ***Note3: Other aspect(s) is not precluded*** * ***Note4: the above study should consider the associated collaboration levels***   Mod: According to the agreement made in agenda 9.2.1, LCM consists of all phase/aspects of AI/ML operations (e.g. data collection, inference, …). Thus, LCM is not in the same level of other bullets.  **vivo’s reply:** If so, we suggest the following wording, which should be safer than the current proposal as we don’t miss anything for now. We can further switch the order of the 3rd and 4th bullets if it would look better.  ***Proposal 2.6.1c: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***   * ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)***    + ***Note1: Online training and/or offline training is a separate discussion*** * ***New or enhanced mechanism(s) to facilitate AI/ML inference*** * ***New or enhanced mechanism(s) to facilitate life cycle management aspects other than data collection, inference and performance monitoring  ~~AI/ML model activation/deactivation/selection/switching and fall-back operation~~*** * ***New or enhanced mechanism(s) to facilitate AI/ML model ~~performance~~ monitoring*** * ***AI-related UE capability and reporting*** * ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting*** * ***Note3: Other aspect(s) is not precluded*** * ***Note4: the above study should consider the associated collaboration levels*** |
| CATT | Share the same view as vivo.  Mod: According to the agreement made in agenda 9.2.1, LCM consists of all phase/aspects of AI/ML operations (e.g. data collection, inference, …). Thus, LCM is not in the same level of other bullets. |
| Panasonic | Support. |
| Xiaomi | Support |
| FUTUREWEI | Support Proposal 2.6.1c. |
| FUJITSU | Support |
| Mod | The agreement for LCM made in agenda 9.2.1 has covered this proposal. |

### Data collection

The proposals/ observations related to the general principles are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 4: Regarding BM-Case1, when Set B is a subset of Set A, study the standards impact to enable gNB to collect assistance attributes that are needed for model training and model inference.* |
| Huawei[2] | *Proposal 9: Study potential specification impact for AI/ML-based beam prediction considering the following aspects:*   * *AI/ML model training procedure* * *Enhancement for RSRP report and beam ID report for inference* |
| BJTU[12] | *Proposal #4: Study potential specification impact for AI/ML-based HSR beam management, considering the following aspects:*   * + *Collaboration procedure between UE and gNB.*   + *AI/ML model deployment, training and inference procedure.* |
| CATT[13] | *Proposal 15: The following spec impact of AI/ML based beam management can be considered:*   * *Signaling/procedure of AI model training/updating/fallback;* * *Interface of AI model, i.e. relationship between measured RS and reported information;* * *New procedure for RS measurement and reporting;* * *Signaling/procedure design on exchanging AI-related/non-AI-related assistance information.* |
| NVIDIA[16] | *Proposal 6: 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.* |
| Ericsson[24] | *Proposal 7 New or enhanced mechanism(s) including CSI-report-based, SRS-based and RRC-message-based frameworks to facilitate NW data collection for beam management use cases should be studied* |

Some tdocs discuss the data collection for AI model training. Based on the proposal, the following proposal is suggested as a starting point for further discussion

###### Proposal 2.6.2 (Closed)

***Proposal 2.6.2: For the data collection for AI/ML model training (if supported), study the following aspects as a starting point:***

* ***Procedure of data collection***
* ***Signaling/configuration for data collection***
* ***Content/type of the collected data***
* ***Reference signals***
* ***Other aspect(s) is not precluded***

***Proposal 2.6.2a: For the data collection for AI/ML model training (if supported), study the following aspects as a starting point:***

* ***Procedure of data collection***
* ***Signaling/configuration for data collection***
* ***Content/type of the collected data***
* ***~~Reference signals~~***
* ***signaling aspects related to assistance information (if supported)***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Support |
| CATT | Support |
| ZTE | Support |
| MediaTek | Support |
| NEC | Support |
| Lenovo | Support |
| CAICT | Support |
| NVIDIA | Support |
| FUTUREWEI | We support the proposal. |
| Google | We think this can be deprioritized. |
| Xiaomi | It can be discussed after the discussion on on-UE training, on-network training, input and output. The spec impact will be different for different training side, input and output. |
| Spreadtrum | Support |
| PANASONIC | Support |
| vivo | OK |
| Sony | Agree with proposal |
| OPPO | Support |
| qualcomm | Support, also suggest adding signaling aspects related to assistance information for data collection, which could be in the form of meta-data. It is important to mention that this assistance information is different from the one discussed in the context of ‘inputs to AI/ML models.’ |
| Fujitsu | support |
| HW/HiSi | OK |
| NTT DOCOMO | Support |
| Ericsson | Support. We think the term “reference signals” are unclear. Propose to remove this. |
| Intel | Not clear what reference signals means |
| Samsung | Support |
| CMCC | Support |
| MOd | The proposal is updated according to the inputs |
| CAICT | Fine with the update |
| Fujitsu | support |
| Lenovo | Fine |
| Xiaomi | Fine with the proposal 2.6.2a |
| NEC | OK for the updated proposal. |
| Samsung | OK. |
| CATT | Support |
| Qualcomm | Support Proposal 2.6.1b. |
| Ericsson | Support |
| Spreadtrum | Support the proposal 2.6.2a |
| Mod | Please see the section of “Summary of discussion”.  The discussion is closed |

### AI/ML inference for BM-Case1 & BM-Case2

#### General/common aspects

There are some contributions discussing the detailed spec impacts of BM-Case1 and BM-Case2. The related proposals/observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 3: Regarding BM-Case1, when Set B is a subset of Set A, study the standards impact to enable gNB to map the received L1-RSRP measurements to the corresponding Tx-Rx beam pairs.*  *Proposal 4: Regarding BM-Case1, when Set B is a subset of Set A, study the standards impact to enable gNB to collect assistance attributes that are needed for model training and model inference.* |
| Fujitsu[7] | *Proposal 2: For the NW-side model, study the following potential specification impacts for spatial- domain DL beam prediction*   * *Signaling to carry information about RX beam pattern.* * *Beam measurement reporting (non-group-based and group-cased) including RX beam information.*   *Proposal 3: For the UE-side model, study the following potential specification impacts for spatial- domain DL beam prediction*   * *Signaling to carry information about TX beam pattern.* * *Signaling to inform UE about the mapping of RSs and TX beams.* * *Signaling to inform NW about the subset of RSs.* |
| NEC[14] | Proposal 15: Study the method of indicating the predicted beams and corresponding beam application/dwelling times. |
| Xiaomi[19] | *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 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.* |
| Samsung[21] | Proposal 2: For BM-Case1, further study the specification impacts for AI/ML inference at NW side considering the following aspects.   * Enhancement on L1 beam report mechanism * Assistance information for beam prediction   Proposal 3: For BM-Case1, further study the specification impacts for AI/ML inference at UE side considering the following aspects.   * Assistance information for AI/ML inference at UE side * Enhancement on L1 beam report mechanism   Proposal 5: For BM-Case2, further study the specification impacts for AI/ML inference at NW side considering the following aspects.   * Enhancement on L1 beam report mechanism   Proposal 6: For BM-Case2, further study the specification impacts for AI/ML inference at UE side considering the following aspects.   * Enhancement on L1 beam report mechanism * UE-side case/events that can leverage the predicted/future L1-RSRP |
| LGE[22] | Proposal #3: Consider UE assistance/reporting for determining Set A.  Proposal #5: For NW-side AI/ML in BM-Case2, consider enhancements on UE reporting and/or beam indication.  Proposal #6: For UE-side AI/ML in BM-Case2, consider enhancements on beam reporting. |
| CMCC[23] | Proposal 2: For model inference of spatial domain beam prediction at gNB side, CSI report framework needs further enhancement.  Proposal 3: For model inference of spatial domain beam prediction at UE side, CSI report framework needs further enhancement. |
| Nokia[25] | Proposal 3: Further study of the DL Tx beam prediction failure detection/recovery procedure and model switching procedure.  Proposal 11: RAN1 to study the impact of data collection on radio link failures and time of outage.  Proposal 16: For the use case of DL Rx beam prediction, UE needs to report its Rx beam capability and the needed Rx beam sweeping number, which may be different from the UE Rx beam capability max Number of Rx Beam. |
| QC[27] | Proposal 1: Study the signalling aspects related to beam blockage/failure prediction, as a sub-use case of temporal beam prediction.  Proposal 3: Study the signalling aspects related to gNB sending assistance information to help UE with data collection for training, for the purpose of temporal beam prediction.  • Examples of such assistance information: information about gNB beam shape, beam boresight directions, 3dB beamwidth, etc., information about gNB antenna array structure.  Proposal 4: Study and evaluate the benefits of temporal beam prediction at UE and gNB and the associated signalling needed to assist or enable beam prediction at each side.   * + The trade-off between beam prediction accuracy and required signalling overhead for UE-side and gNB-side inference should be considered in the study.     - UE-side inference:       * Study enhanced UE L1 report including information from temporal beam prediction       * Study signalling aspects related to assistance information from gNB to help beam prediction at UE     - gNB-side inference:       * Study enhanced UE L1 report to improve beam prediction quality at gNB   Proposal 8: For UE-side training, and for the agreed sub-use cases (Alt. 1 and Alt. 2) study the signalling aspects related to gNB sending assistance information to help UE with data collection for training, for the purpose of spatial domain beam prediction.  Proposal 9: RAN1 should study and evaluate the benefits of spatial (+time) domain beam prediction at UE and gNB and the associated signalling needed to assist or enable beam prediction at each side.   * + The trade-off between beam prediction accuracy and required signalling overhead should be considered in the study.   + UE-side inference:     - Study enhanced UE L1 report including information from spatial domain beam prediction     - Study signalling aspects related to assistance information from gNB to help beam prediction at UE   + gNB-side inference:     - Study enhanced UE L1 report to improve beam prediction quality at gNB |
| Apple[28] | Proposal 3: Study spatial domain beam prediction with measurement for limited number of beams as well as a flexible beam measurement and report framework to support dynamic activation/deactivation of beam measurement reference signal and beam report.  Proposal 4: Study time domain beam prediction based on past measurement results as well as TCI activation/indication to facilitate the beam prediction in time domain. |
| DCM[29] | Observation 1: Enhancements on beam selection policy in CSI reports might be potential specification impacts for spatial domain beam estimation.  Proposal 4: CSI report should be enhanced to improve the performance of time-domain beam prediction, if time-domain beam prediction is supported as sub use-case.  Proposal 7: Support mechanisms to provide DL Tx beam information from NW to UE for DL beam prediction with UE side model, if it is beneficial for the beam prediction with UE side model. |

This section focuses on the common issues of the sub use cases. Some dedicated spec impact or more detailed impacts for some use cases will be discussed in the subsequent sections.

Proposal 2.6.3.1

***Proposal 2.6.3.1: In order to facilitate the AI/ML model inference, study the following aspects as a starting point:***

* ***Enhanced or new beam measurement and/or beam reporting***
* ***Beam indication of the predicted beam(s)***
* ***Enhanced or new signaling for measurement configuration/triggering***
* ***Signaling of assistance information (if supported)***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Support |
| CATT | Support |
| ZTE | Support |
| MediaTek | Support |
| NEC | Support |
| Lenovo | Support |
| CAICT | Support |
| NVIDIA | Support |
| FUTUREWEI | Support |
| Google | Support |
| Spreadtrum | Support |
| PANASONIC | Support |
| vivo | Suggest to add Reference signal as one aspect for study. For example, either BM Case 1 or Case 2 can facilitate overhead reduction of RS (in spatial domain or time domain). It is not just related with data collection or measurement, but also related with saving RS overhead to have more resource for data channel scheduling.  Hence we think it is needed to add the following bullet  ***Reference signal (e.g., overhead reduction)***  Mod: It seems not parallel to the listed bullet. Please feel free to correct me if I misunderstand something. |
| Sony | It maybe better to give a list of assistance information such as location, UE moving direction  Mod: AI 9.2.3.1 may have some discussion on assistance information. Thus, we need to avoid the duplicated discussion. Moreover, it is the next level details |
| OPPO | Support |
| Qualcomm | Support |
| Fujitsu | support |
| HW/HiSi | Support |
| NTT DOCOMO | Support the proposal. |
| Ericsson | Support |
| Intel | OK |
| Samsung | Support |
| CMCC | Support |
| vivo | Thanks FL for the reply. If the intention is that RS aspect has been covered by some of the bullets above, it is better to make it more clear. We suggest the following.  ***Proposal 2.6.3.1: In order to facilitate the AI/ML model inference, study the following aspects as a starting point:***   * ***Enhanced or new beam measurement and/or beam reporting, including RS overhead reduction*** * ***Beam indication of the predicted beam(s)*** * ***Enhanced or new signaling for measurement configuration/triggering*** * ***Signaling of assistance information (if supported)*** * ***Other aspect(s) is not precluded*** |
| Qualcomm | For the first bullet, given the signaling of assistance information is still ongoing, we do not want to limit the scope of ‘enhanced UE report’ to ‘beam’ measurement and ‘beam’ reporting. So, we suggest the following change for first bullet:   * ***Enhanced or new UE report, e.g., enhanced or new beam measurement and/or beam reporting*** |

Proposal 2.6.3.1a

The proposal is updated by combing the suggestion from vivo and QC

***Proposal 2.6.3.1a: In order to facilitate the AI/ML model inference, study the following aspects as a starting point:***

* ***Enhanced or new UE report, e.g., Enhanced or new beam measurement and/or beam reporting, including RS overhead reduction***
* ***Beam indication of the predicted beam(s)***
* ***Enhanced or new signaling for measurement configuration/triggering***
* ***Signaling of assistance information (if supported)***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
| NTT DOCOMO | Support the proposal. |
| NEC | Support the updated proposal |
| vivo | Generally fine. Just a small change on the first sub-bullet as the example talks about not just reporting but also measurement.  ***Proposal 2.6.3.1a: In order to facilitate the AI/ML model inference, study the following aspects as a starting point:***   * ***Enhanced or new UE report/measurement, e.g., Enhanced or new beam measurement and/or beam reporting, including RS overhead reduction*** * ***Beam indication of the predicted beam(s)*** * ***Enhanced or new signaling for measurement configuration/triggering*** * ***Signaling of assistance information (if supported)*** * ***Other aspect(s) is not precluded*** |
| CATT | Agree with vivo’s updated. |
| Panasonic | Support. |
| Xiaomi | Support the proposal |
| Fujitsu | support |
| Charter | More study required on UE capability for enhanced beam reporting  Mod: UE capability can be discussed in Section 2.6.5 |

###### Proposal 2.6.3.1b (H)

Proposal 2.6.3.1b is updated from 2.6.3.1a by adding “measurement” as suggested by vivo and CATT.

***Proposal 2.6.3.1b: In order to facilitate the AI/ML model inference, study the following aspects as a starting point:***

* ***Enhanced or new UE report/measurement, e.g., Enhanced or new beam measurement and/or beam reporting, including RS overhead reduction***
* ***Beam indication of the predicted beam(s)***
* ***Enhanced or new signaling for measurement configuration/triggering***
* ***Signaling of assistance information (if supported)***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
| Futurewei | We are ok with the proposal in general. However, this proposal is for “***facilitate the AI/ML model inference”,*** we think the wording of “***including RS overhead reduction”*** is confusing in this proposal, thus, we suggest removing it. |
| ZTE | The first sub bullet seems a little confusing. We suggest the following revision.   * Enhanced or new resource/reporting configuration/triggering/activation and UE measurement |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

#### AL/ML inference at UE side (BM-Case1)

On top of Proposal 2.6.3.1, more details or new aspect will be added based on more inputs. The following proposal is a skeleton and more inputs are expected.

###### Proposal 2.6.3.2 (Low priority)

***Proposal 2.6.3.2: When AL/ML inference is carried out at UE side (UE-side model), study the following aspects as a starting point:***

* ***Signaling of the relationship between Set A and Set B***
* ***…***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | Support |
| CATT | At this stage, we don’t prefer to discuss the detail procedure at UE can gNB side separately, since the various simulation results will be discussed in EVM agenda. Proposal 2.6.3.1 is enough. |
| ZTE | We prefer not to discuss further spec details until the discussion on AI input/output is sufficiently stable. |
| mediatek | Agree with ZTE that AI/ML inputs and outputs need to be agreed first to continue further discussions on this aspect. |
| NEC | Agree with CATT. It is too early to discuss this proposal. |
| Lenovo | Support |
| CAICT | We are open to discuss the details and this proposal could be discussed later. |
| NVIDIA | Incomplete proposal. |
| FUTUREWEI | We think the details regarding signaling of the relationship between Set A and Set B can be discussed later in the study while potential standards impacts associated with UE-side mode (and NW-side model) should be identified and discussed first. |
| Google | We are not sure whether it is feasible to do it in UE side. More study could be needed. |
| Spreadtrum | Agree with ZTE. |
| Panasonic | We support the proposal as UE needs to know the relationship between Set A and Set B for UE side model but not for NW side model. |
| vivo | This can be discussed after or together with proposal 2.6.3.1. This is a next level of details for 2.6.3.1. |
| HW/HiSi | Neutral |
| NTT DOCOMO | Support the proposal. |
| Ericsson | We don’t understand the proposal. It should rather be part of potential assistance information discussion |
| Intel | Proposal is unclear |
| Samsung | Open to further study this aspect but the discussion could be at later stage. |
| CMCC | We also think whether to signal the relationship between Set A and Set B can be discussed later. Moreover, the concept of “relationship” is not clear. |
| Charter | It is too early to discuss implications of AI/ML inference at UE and should be pushed to a later study. |

#### AL/ML inference at gNB side (BM-Case1)

On top of Proposal 2.6.3.1, more details or new aspect will be added based on more inputs.

**Moderator recommendation**: TBD

|  |  |
| --- | --- |
| Company | Comments |
| CATT | See above. We think currently Proposal 2.6.3.1 is enough. |
| MediaTek | We think that AI/ML inputs and outputs need to be agreed first to continue further discussions on this aspect. |
| NEC | As mentioned before. It is too early to discuss this proposal. |
| vivo | This can be discussed after or together with proposal 2.6.3.1. This is a next level of details for 2.6.3.1. |
| Fujitsu | In this stage, it’s not necessary to separately discuss the specification impacts for UE-side and NW-side model. |

#### AL/ML inference at UE side (BM-Case2)

On top of Proposal 2.6.3.1, more details or new aspect will be added based on more inputs. The following proposal is a skeleton and more inputs are expected.

###### Proposal 2.6.3.4 (Low priority)

***Proposal 2.6.3.4: When AL/ML inference is carried out at UE side (UE-side model), study the following aspects as a starting point:***

* ***Signaling of the relationship between Set A and Set B***
* ***Beam reporting enhancement, e.g.,***
  + ***associated timing information of each measurement result (explicit or implicit)***
  + ***reported measurements for a larger number of beams***
* ***…***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | First bullet can be removed if Alt3(SetA=SetB) can be agreed for BM-Case2. |
| CATT | See above. We think currently Proposal 2.6.3.1 is enough. |
| MediaTek | We are supportive to the proposal in general. However, as mentioned before, we think that AI/ML inputs and outputs need to be agreed first to continue further discussions on this aspect. |
| NEC | As mentioned before. It is too early to discuss this proposal. |
| Lenovo | Support |
| CAICT | We are open to discuss the details and this proposal could be discussed later. |
| NVIDIA | Incomplete proposal. |
| Google | We are not sure whether it is feasible to do it in UE side. More study could be needed. |
| Panasonic | Support. We agree with LGE that the first bullet applies to Alt1 and Alt2. |
| vivo | This can be discussed after or together with proposal 2.6.3.1. This is a next level of details for 2.6.3.1. |
| Qualcomm | Support the first bullet. Agree with second bullet in general, however details of beam reporting enhancements should be discussed later. |
| HW/HiSi | Postpone this discussion until we have progressed further with the basics. |
| NTT DOCOMO | We are not sure why the number of reported beams would be increased in BM-Case 2 with UE side model. Is it because one beam reporting instance includes beam measurements at multiple time instances in future? |
| Ericsson | We think it should list one item, and the reporting enhancement could be a starting point. |
| Intel | We should come back to this after more progress on BM-Case 2 details |
| Samsung | Open to further study this aspect but the discussion could be at later stage. |
| CMCC | Same as the comment of Proposal 2.6.3.2. |

#### AL/ML inference at gNB side (BM-Case2)

On top of Proposal 2.6.3.1, more details or new aspect will be added based on more inputs.

**Moderator recommendation**: TBD

|  |  |
| --- | --- |
| Company | Comments |
| CATT | We think currently Proposal 2.6.3.1 is enough. |
| vivo | This can be discussed after or together with proposal 2.6.3.1. This is a next level of details for 2.6.3.1. |
|  |  |

### Life cycle management

There are many contributions discussing potential spec impacts of the life cycle management of AI/ML model(s). The related proposals/ observations are copied as below:

|  |  |
| --- | --- |
| FUTUREWEI[1] | *Proposal 5: Regarding AI/ML-based beam management, study the standards impact, including AI/ML related UE capability reporting, related to AI/ML model selection/configuration (like activation/deactivation) in case multiple trained AI/ML models are deployed.* |
| Vivo[4] | Proposal 18: 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 19: At least the following life cycle management component need to be studied for beam management: model activation, data collection for model inference, model inference, data collection for model monitoring, model monitoring and model deactivation.  Proposal 22: Study specification impact of model performance monitoring for both spatial domain and temporal domain beam prediction regarding at the following aspects:  a) Monitoring configuration and/or activation conditions  b) Monitoring resources  c) Monitoring metrics  d) Monitored results reporting  e) Impairments for monitoring, e.g., how to monitor with non-ideal labels |
| Google[9] | Proposal 4: 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 8: 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 9: 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. |
| OPPO[11] | Proposal 11: Study the performance monitoring mechanism of AI/ML model for beam prediction. |
| NEC[14] | Proposal 7: Study the mechanism of model update, e.g., fine-tuning.  Proposal 9: Study the mechanism of model selection.  Proposal 13: Study the direct or indirect mechanisms on evaluating the performance of model inference. |
| Lenovo[15] | Proposal 8: Dynamic switching between AI/ML based beam prediction and non-AI/ML based beam report should be supported. |
| Xiaomi[19] | *Proposal 10: gNB to transmit all beams in set A periodically/semi-persistently/ a-periodically for performance monitoring.*  *Proposal 11: Threshold of beam prediction accuracy related KPIs can be used for performance monitoring.*  *Proposal 12: Study the mechanism for AI model update/disable/deactivation request from UE.*  *Proposal 13: Study the mechanism for AI model disable/deactivation/update by gNB.* |
| CMCC[23] | Proposal 4: For model monitoring of spatial domain beam prediction, model monitoring performance metric needs to be determined, the ignaling for obtaining/reporting model monitoring performance metric and indicating/requesting model updating/switching/fallback needs further enhancement. |
| Ericsson[24] | Proposal 10 Study mechanisms for performance monitoring for beam prediction AI/ML models  Proposal 11 Study mechanisms to activate/deactivate beam prediction AI/ML models, and potential fallback mechanisms |
| QC[27] | Proposal 5: For temporal beam prediction, study the signalling aspects related to exchanging information about beam prediction quality and a metric for beam prediction quality   * Study the impact of beam prediction quality on activating/deactivating AI/ML module at UE.   Proposal 6: For temporal beam prediction, 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. |
|  | Proposal 10: For spatial domain beam prediction, RAN1 should study the signalling aspects related to exchanging information about beam prediction quality and a metric for beam prediction quality   * Study the impact of beam prediction quality on activating/deactivating AI/ML module at UE   Proposal 11: For spatial domain beam prediction, RAN1 should 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 additional reference signals. |
| DCM[29] | Proposal 5: Beam measurement of Set A for model performance monitoring should be studied as potential specification impacts.  Proposal 6: Study NW-based model monitoring and UE-based model monitoring in beam prediction with UE-side model. |
| Panasonic[30] | Proposal 3: For AI/ML inference at UE side, study methods for AI/ML model configuration, activation and monitoring. |

Some of the proposals are focusing on the high-level functionality of LCM, whereas some proposals are focusing on the detailed design, e.g., which beam is used for the reference of performance monitoring. Generally speaking, all the proposals belong to one of the following aspects:

* AI/ML Model management
* Update of AI/ML model
* Performance monitoring

Thus, moderator suggest the following proposal as a starting point, which focus on the high-level aspects of potential spec impacts. More details can be discussed latter.

###### Proposal 2.6.4-1 (closed)

***Proposal 2.6.4-1: For AI model life cycle management of BM-Case1 and BM-Case2, support to investigate specification impacts from the following aspects***

* ***Mechanisms for AI/ML model configuration/activation/deactivation/selection/switching and fall-back operation***
* ***Mechanisms for AI model re-tuning***
* ***Mechanisms for performance monitoring***
* ***Other aspect(s) is not precluded***

***Proposal 2.6.4-1a: For AI model life cycle management of BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***

* ***Mechanisms for AI/ML model configuration/activation/deactivation/selection/switching and fall-back operation***
* ***~~Mechanisms for AI model re-tuning~~***
* ***Mechanisms for performance monitoring***
* ***Other aspect(s) is not precluded***
* ***Note: the above study should consider the associated collaboration levels***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | What is difference of the second bullet from the first bullet? If there is no difference, the second bullet can be removed. |
| CATT | We wonder what the spec impacts on AI model re-tuning are. Except for the first bullet and data collection, the AI model re-tuning is a kind of implementation. |
| ZTE | We prefer to remove the second bullet. |
| MediaTek | Further discussions are needed to include AI/ML model re-tuning (second bullet) |
| NEC | Support |
| Lenovo | Prefer to include the second bullet in the first bullet. |
| CAICT | The second bullet needs further clarification. |
| NVIDIA | The second bullet needs clarification. |
| Futurewei | We think the 2nd bullet is not needed as the specification impact for retuning is covered by data collection which is covered by ***Proposal 2.6.1 and Proposal 2.6.2,*** and the first bullet. |
| Google | Support |
| Xiaomi | Further clarification is needed for the second bullet. |
| Spreadtrum | The second bullet should be more clearly defined |
| PANASONIC | Support |
| vivo | One critical aspect in this discussion is the corresponding collaboration levels. Different levels may require different procedures. Hence we think it is needed to add the following bullet.  ***Associated collaboration levels*** |
| Sony | Support the model switching based on the different applications of applications as the propagation environment is quite different as also mentioned in OPPO [11]. Performance degradation will be experienced when the models are used in different environment from the training environment. |
| OPPO | Support |
| qualcomm | For second bullet, the terminology needs to be elaborated, and this can also follow the discussions in general framework agenda item. |
| Fujitsu | The definition of re-tuning is not clear. Prefer to wait the definition discussion in 9.2.1 |
| HW/HiSi | The second bullet “re-tuning” could be included in the examples given of the first bullet. |
| NTT DOCOMO | Support the proposal. Since the complexity of AI model updating is higher than model activation/deactivation, we prefer to keep the first bullet and second bullet separate as the current proposal. |
| Ericsson | Support bullet 1 and 3. Second bullet should be removed. What is meant by re-tuning? |
| Intel | Difference between bullet 2 and 3 is unclear |
| Samsung | The second bullet requires more clarification. In terms of AI model re-tuning, is it the same as ‘online’ refinement/re-tuning of AI model parameters?  If so, we suggest to remove the second bullet.  Also, we suggest the similar revision as proposal 2.6.3.1.  ***Proposal 2.6.4-1: For AI model life cycle management of BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects.*** |
| CMCC | Prefer to incorporate the first and second bullet. |
| MOd | The proposal is updated   * remove bullet 2 * Add note 1 to address vivo’s concern |
| Mod2 | It is merged to Proposal 2.6.1 |
| vivo | We don’t agree to remove “AI model retuning”. We think it is a useful tool to ensure AI generalization performance under different channel environments and cell/UE configurations. We don’t agree to remove it without study. |
| Lenovo | We agree with vivo that “AI model retuning” should not be removed. |
| NEC | We are open whether “AI model retuning” is removed or not. And if it is not removed, we prefer to incorporate the first and second bullet. |
| Qualcomm | Support Proposal 2.6.4-1a. |
| HW/HiSi | Neutral, but since model configuration is not confirmed in 9.2.1, we suggest to set the first bullet in brackets and re-write this proposal as:  ***Proposal 2.6.4-1a: For AI model life cycle management of BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***   * ***[Mechanisms for AI/(ML model configuration)/activation/deactivation/selection/switching and fall-back operation]*** * ***~~Mechanisms for AI model re-tuning~~*** * ***Mechanisms for performance monitoring*** * ***Other aspect(s) is not precluded*** * ***Note: the above study should consider the associated collaboration levels*** |
| Ericsson | Support |
| Futurewei | We are ok with the updated proposal. We think the model retuning can be covered by the standards impact for training and the other bullet items in this proposal. |
| Charter | Support |
| Mod | The agreement for LCM made in agenda 9.2.1 has covered this proposal. |

Proposal 2.6.4-2

***Proposal 2.6.4-2: Regarding the performance monitoring for BM-Case1 and BM-Case2, support to investigate specification impacts from the following aspects***

* ***Performance metric(s)***
* ***Benchmark/reference for the performance comparison***
* ***Signalling/procedure for information collection***
* ***Other aspect(s) is not precluded***

***Proposal 2.6.4-2a: Regarding the performance monitoring for BM-Case1 and BM-Case2, support to investigate specification impacts from the following aspects***

* ***Performance metric(s)***
* ***Benchmark/reference for the performance comparison***
* ***Signalling/procedure for information collection***
* ***assistance signaling (e.g., auxiliary reference signals)***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
| LGE | The first two bullets may belong to EVM agenda?  Mod: In my understanding, in the agenda, we can discuss the aspects/mechanism for a feature. The detailed design should consider the output of evaluation results. Thus, from my perspective, there is no conflict between these two agendas. |
| CATT | We think at current stage Proposal 2.6.4-1 is enough. We may first discuss the performance for BM-Case1 and BM-Case2 in EVM agenda. The details for performance monitoring can be discussed further.  Mod: I understand that the final design will depends on the evaluation results. On the other hand, the intention of this proposal is to suggest companies to focus on the study/evaluation on some typical aspects and facilitate the further discussion. I failed to say any conflict between these two agendas if we agree this proposal. If I missed something, please fell free to correct me. |
| ZTE | Support the FL’s proposal. |
| MediaTek | We support this proposal. |
| NEC | Support |
| Lenovo | Support |
| CAICT | Support |
| NVIDIA | Support |
| Futurewei | We agree with CATT that Proposal 2.6.4-1 already covers specification impact related to performance monitoring.  Mod: Please see the reply to CATT |
| Google | Support |
| Xiaomi | Support |
| Spreadtrum | Support |
| PANASONIC | We share similar view as LGE and CATT.  Mod: Please see the reply to CATT |
| vivo | We are generally okay to list the detailed aspects for performance monitoring. |
| Sony | Before performance evaluation we shall study the availability of test data  Mod: Please see the reply to CATT |
| OPPO | Support |
| Qualcomm | Support the proposal but suggest adding ‘assistance signaling’ which could be in the form of auxiliary reference signals from gNB to help UE in comparing predicted measurements to actual ones and see how AI/ML model is doing. The ‘information collection’ mentioned in third bullet is not quite clear with regards to performance monitoring and needs some elaboration.  Mod: For example, if a metric is agreed, then there may be some signaling to collect the information regarding the quality in terms of the metric.  “assistance signaling” is added |
| Fujitsu | support |
| HW/HiSi | Support |
| NTT DOCOMO | Support the proposal. |
| Ericsson | Support |
| Intel | OK |
| Samsung | Same view as CATT.  Mod: Please see the reply to CATT |
| CMCC | Support |
| Mod | The Proposal is updated to Proposal 2.6.4-2a |
| CAICT | Fine with the update |
| Fujitsu | support |
| Lenovo | Support |
| NEC | Support the updated proposal. |
| Qualcomm | Support updated Proposal 2.6.4-2. |
| Ericsson | We think it is sufficient with “assistance signaling”. “Auxiliary reference signals are unclear to us”. It would rather comprise auxiliary measurements. Propose to remove it for now. |
| Futurewei | We are ok with proposal 2.6.4-2a. |

###### Proposal 2.6.4-2b (closed)

The agreement made in AI 9.2.1 uses the terminology “model monitoring” rather than “performance monitoring”. Thus, the proposal is updated to keep the consistency on the terminology based on proposal 2.6.4-2a. The example is also removed as suggested by Ericson.

***Proposal 2.6.4-2b: Regarding the ~~performance~~ model monitoring for BM-Case1 and BM-Case2, support to investigate specification impacts from the following aspects***

* ***Performance metric(s)***
* ***Benchmark/reference for the performance comparison***
* ***Signalling/procedure for information collection related model monitoring***
* ***assistance signaling ~~(e.g., auxiliary reference signals)~~***
* ***Other aspect(s) is not precluded***

|  |  |
| --- | --- |
| Company | Comments |
| Futurewei | We are ok with the updates. |
| NTT DOCOMO | Support the proposal. |
| NEC | We note that the following are mentioned in Proposal 2.6.3.1a:   * ***Enhanced or new UE report, e.g., Enhanced or new beam measurement and/or beam reporting, including RS overhead reduction***   Generally speaking, model monitoring will involve model inference. For example, the measured results needs to be compared with the predicted results from model inference. So, we think it is necessary to add a consistent alternative in model monitoring.  Mod: In the agreement made in agenda 9.2.1, model monitoring and model inference operation are two parallel aspects. Thus, even if model monitoring will need to use some output of the AI/ML inference, we can discuss the spec impact separately and we don’t need to couple the study aspects of inference and monitoring. |
| CATT | Fine with the updated. |
| Xiaomi | We suggest to add “model monitoring results report/indication”. We think it is not covered by “***Signalling/procedure for information collection related model monitoring***”. |
| Fujitsu | support |
| Charter | Support |
| Mod | The corresponding agreement is made in online session |

### Capability

The related proposals/ observations are copied as below:

|  |  |
| --- | --- |
| Lenovo[15] | *Proposal 4: Study UE/NW capability related signaling corresponding to AI-based beam management under different network-UE collaboration levels.* |
| NVIDIA[16] | *Proposal 11: 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.* |

**Moderator recommendation**: TBD

# Summary of Discussion

## Offline agreement for Wednesday online session (closed)

***Proposal 2.2.2-1a: For the sub use case BM-Case1, support the following alternatives for further study:***

* ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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 narrow and wide beam terminology is for SI discussion only and have no specification impact~~***
* ***Note2~~3~~: The ~~codebook constructions~~ beam patterns of Set A and Set B can be clarified by the companies.***

According to the inputs received so far, all companies are fine with Proposal 2.2.2-1a.

***Proposal 2.6.2a: For the data collection for AI/ML model training (if supported), study the following aspects as a starting point:***

* ***Procedure of data collection***
* ***Signaling/configuration for data collection***
* ***Content/type of the collected data***
* ***~~Reference signals~~***
* ***signaling aspects related to assistance information (if supported)***
* ***Other aspect(s) is not precluded***

According to the inputs received so far, all companies are fine with Proposal 2.6.2a

## Proposals for Thursday online session

***Proposal 2.2.2-2a: For the sub use case BM-Case2, support the following alternatives for further study:***

* ***Alt.1: Set A and Set B are different (~~e.g. Set A consists of narrow beams and Set B consists of wide beams~~ 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: Predicted beam(s) are selected from Set A and measured beams used as input are selected from Set B.***
* ***~~Note2: The narrow and wide beam terminology is for SI discussion only and have no specification impact~~***
* ***Note2~~3~~: The ~~codebook constructions~~ beam pattern of Set A and Set B can be clarified by the companies.***

We made similar agreement for BM-Case2 yesterday. Proposal 2.2.2-2a is a counterpart for BM-Case2. Based on the inputs received so far, it seems acceptable to all companies.

***Proposal 2.6.4-2b: Regarding the ~~performance~~ model monitoring for BM-Case1 and BM-Case2, support to investigate specification impacts from the following aspects***

* ***Performance metric(s)***
* ***Benchmark/reference for the performance comparison***
* ***Signalling/procedure for information collection related model monitoring***
* ***assistance signaling ~~(e.g., auxiliary reference signals)~~***
* ***Other aspect(s) is not precluded***

Based on the inputs received so far, it seems acceptable to majority companies.

***Proposal 2.6.3.1b: In order to facilitate the AI/ML model inference, study the following aspects as a starting point:***

* ***Enhanced or new UE report/measurement, e.g., Enhanced or new beam measurement and/or beam reporting, including RS overhead reduction***
* ***Beam indication of the predicted beam(s)***
* ***Enhanced or new signaling for measurement configuration/triggering***
* ***Signaling of assistance information (if supported)***
* ***Other aspect(s) is not precluded***

Proposal 2.6.3.1b is updated from 2.6.3.1a by adding “measurement”.

***Proposal 2.4b: Regarding the sub use case BM-Case1 and BM-Case2, support to 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 ~~Top-N~~ DL Tx and/or Rx beams*** 
  + ***FFS: how to select the N ~~Top-N1~~ 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)***)
* ***Alt.2: Tx and/or Rx Beam ID(s) of the N predicted ~~Top-N~~ 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)***
* ***Alt.3: Tx and/or Rx Beam angle(s) and the predicted L1-RSRP (optional) of the N predicted ~~Top-N~~ DL Tx and/or Rx beams***
* ***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”~~***

Proposal 2.4b is updated from Proposal 2.4a by removing the “Top-N” as suggested by vivo. Since “Top-N” is removed, Note 6 becomes unnecessary. Thus, Note 6 is also removed

***Proposal 2.6.1c: For the sub use case BM-Case1 and BM-Case2, support to investigate the necessity and/or specification impacts from the following aspects***

* ***New or enhanced mechanism(s) to facilitate data collection for UE/NW model operations (e.g., training)*** 
  + ***Note1: Online training and/or offline training is a separate discussion***
* ***New or enhanced mechanism(s) to facilitate AI/ML inference***
* ***New or enhanced mechanism(s) to facilitate AI/ML model activation/deactivation/selection/switching and fall-back operation***
* ***New or enhanced mechanism(s) to facilitate AI/ML model ~~performance~~ monitoring***
* ***AI-related UE capability and reporting***
* ***Note2: mechanism(s) may include procedure, signaling, reference signal, reporting***
* ***Note3: Other aspect(s) is not precluded***
* ***Note4: the above study should consider the associated collaboration levels***

# Reference

1. R1-2205754 Continued discussion on other aspects of AI/ML for beam management FUTUREWEI
2. R1-2205893 Discussion on AI/ML for beam management Huawei, HiSilicon
3. R1-2205968 Discussions on Sub-Use Cases in AI/ML for Beam Management TCL Communication
4. R1-2206035 Other aspects on AI/ML for beam management vivo
5. R1-2206071 Discussion on other aspects for AI beam management ZTE
6. R1-2206115 Considerations on AI/ML for beam management Sony
7. R1-2206167 Sub use cases and specification impact on AI/ML for beam management Fujitsu
8. R1-2206182 Discussion for other aspects on AI/ML for beam management InterDigital, Inc.
9. R1-2206198 On Enhancement of AI/ML based Beam Management Google
10. R1-2206251 Other aspects on AI/ML for beam management Rakuten Mobile, Inc
11. R1-2206318 Other aspects of AI/ML for beam management OPPO
12. R1-2206332 Beam management with AI/ML in high-speed railway scenarios BJTU
13. R1-2206394 Other aspects on AI/ML for beam management CATT
14. R1-2206472 Discussion on AI/ML for beam mangement NEC
15. R1-2206513 Further aspects of AI/ML for beam management Lenovo
16. R1-2206523 AI and ML for beam management NVIDIA
17. R1-2206581 Use-cases and specification for beam management Intel Corporation
18. R1-2206606 Discussion on other aspects on AIML for beam management Spreadtrum Communications
19. R1-2206638 Discussion on other aspects on AI/ML for beam management Xiaomi
20. R1-2206678 Discussions on AI-ML for Beam management CAICT
21. R1-2206823 Representative sub use cases for beam management Samsung
22. R1-2206877 Other aspects on AI/ML for beam management LG Electronics
23. R1-2206905 Discussion on other aspects on AI/ML for beam management CMCC
24. R1-2206940 Discussion on AI/ML for beam management Ericsson
25. R1-2206971 Other aspects on ML for beam management Nokia, Nokia Shanghai Bell
26. R1-2206991 Other aspects on AI/ML for beam management MediaTek Inc.
27. R1-2207227 Other aspects on AI/ML for beam management Qualcomm Incorporated
28. R1-2207331 Other aspects on AI/ML for beam management Apple
29. R1-2207404 Discussion on other aspects on AI/ML for beam management NTT DOCOMO, INC.
30. R1-2207506 Discussion on sub use cases of AI/ML beam management Panasonic
31. R1-2207551 Discussion on Performance Related Aspects of Codebook Enhancement with AI/ML Charter Communications, Inc
32. R1-2207590 Discussion on other aspects on AI/ML for beam management KT Corp.
33. R1-2205454 Discussion summary#4 for other aspects on AI/ML for beam management Moderator (OPPO)

# 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 Kang  Haewook Park | [jw.kang@lge.com](mailto:jw.kang@lge.com)  [haewook.park@lge.com](mailto:haewook.park@lge.com) |
| Panasonic | Quan Kuang | quan.kuang@eu.panasonic.com |
| Ericsson | Henrik Ryden | Henrik.a.ryden@ericsson.com |
| Nokia, NSB | Keeth Jayasinghe  Mihai 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 Kyung | gyubum.kyung@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. Ionescu  Samer Henry | dumitru.ionescu@charter.com  C-Samer.Henry@charter.com |
|  |  |  |

# Appendix B: Previous Agreements

## RAN1#110

## 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.