EGUSPHERE-2026-1030

Automated Detection of Low-altitude Isolated Mesospheric Radar Echoes Using YOLOv8: Evidence for a C-Layer Phenomenon near 60 km Altitude?

Yadu Krishnan Krishnakumar, Toralf Renkwitz, and Andreas Ahrens

Answers to Reviewer #1

We are grateful for the detailed and constructive review of our manuscript. Please find below our answers to the remaining points.

It is also peculiar that the authors reference the hypothetical "C-layer" in the title but do mention this layer at all in the article.

Response: That is indeed true and been intentional, but perhaps not advantageous. We so far only used this term in the title to refer to previous publications and make the article more visible. We're far from defining or justifying the term C layer, though we speculate it might be a correct designation. An alternative might also be "lower D layer", similar to F1 in the F-region.

Changes in manuscript: We have widely rewritten the manuscript to also discuss more on the term C layer and the apparent relevance to the observed radar echoes.

Issues to address

Comment 1: In the abstract the authors state "this study might be the first evidence of such a phenomenon through consistent radar observations." While it is true that these results represent the first long-term measurements of layered phenomena around 60 km, layers around this altitude have been reported previously by other authors. For instance, Hocking, W.K. and Vincent, R.A., 1982. Comparative observations of D region HF partial reflections at 2 and 6 MHz. Journal of Geophysical Research: Space Physics, 87(A9), pp.7615-7624, shows persistent layers at 62 km at 2 MHz (Figure 2). Furthermore, Holdsworth, D.A. and Reid, I.M., 1997. An investigation of biases in the full correlation analysis technique. Advances in Space Research, 20(6), pp.1269-1272, show a five-hour averaged SNR profile with a peak at around 57 km (Figure 4). There may well be other such example from early MF radar observations made in Saskatoon, Canada and Christchurch, New Zealand. I suggest the authors conduct a thorough search and acknowledge such observations in their article.

Response: We admit that we have been reticent with other MF/HF radar references, but this is primarily because we are not convinced that we can identify the same radar echoes in these manuscripts. However, we should have supplemented the introduction with additional examples of radar observations at an altitude of approximately 60 km.

Changes in manuscript: We added more explanations to the general formation of the D region, EPP-related events including corresponding references from multiple radar instruments. We also modified the discussion and partly conclusion to work out the difference between the individual types of radar echoes more clearly.

Comment 2: Lines 328 and 358, The authors need to provide more evidence regarding their claim that "GCR represent a significant ionisation source for altitudes just below 60 km during solar minimum years". This conclusion seems to be made solely on the results of Figure 14. While there appears to be an inverse correlation between sunspot number (i.e. solar cycle) and the detection of LIME, and there is a well-established inverse correlation sunspot number and the detection of GCRs, this does not imply GCRs are responsible for LIME. Correlation is not the same as causation. Having said that, I suspect the authors are likely correct, but the evidence presented is somewhat circumstantial. I strongly suggest the authors introduce the word "may" into the sentences on lines 328 and 358 – e.g. "which suggests GCR may represent a significant…"

Response: We agree with the reviewer, the link of LIME to the GCR flux was speculative due to the weak correlation with solar flux.

Changes in manuscript: We have now added GCR detection/occurrence rates to Fig. 12 (in the revised version). For 2024, the reduction of GCR during solar active periods is very prominent including quasi-simultaneous absence of LIME. We also adapted the discussion accordingly.

Comment 3: The authors frequently use the term "bottomside of the ionosphere" to describe the altitude region of interest. This is a misuse of the term, which is typically used to describe the entire ionosphere below the peak of the F-region electron density (hmF2). I strongly recommend the authors replace "bottomside of the ionosphere" with something like "lower ionosphere" or maybe "lower D-region".

Response: We thank for this clarification. With the statement we indeed did not mean the entire D, E, F regions.

Changes in manuscript: We have rephrased this to:

"This lowermost part of the ionosphere (D region and possibly below) …" etc.

Comment 4: Line 69, "4-bit-complementary codes are frequently used to increase the average power". This is incorrect. Complementary codes increase signal-to-noise (SNR) ratio and hence detectability, but they do not increase signal power.

Response: Well, both is correct. On the transmit side, e.g. 4 times the power is used instead of a single pulse of the same range resolution, so the average power is enhanced. For reception, it's correct, the detectability is improved as the random noise or non-matching signals are suppressed.

Changes in manuscript: We've added this to the description. Line 130: "… e.g. 4-bit-complementary codes are frequently used to increase the average power, suppress interfering signals and thus improve the signal-to-noise ratio."

Comment 5: Line 100. What is the mean average precision index, and what do the numbers following mAP mean?

Response: Thank you for this comment. We agree that these evaluation metrics require clarification for readers outside the machine learning domain. A brief explanation has been added to the manuscript at this point.

Changes in manuscript: The following text has been added to the manuscript at Line 172:

"Here, mAP (mean Average Precision) summarises detection accuracy by computing the area under the precision-recall curve, where precision measures the fraction of model detections that are correct, while recall measures the fraction of all true targets in the dataset that the model successfully detected. The numerical suffix denotes the Intersection over Union (IoU) threshold, the minimum spatial overlap required between a predicted and ground-truth bounding box for a detection to be counted as correct."

Comment 6: Line 101. What is an "anchor"? Why is it used?

Response: We agree that the term "anchor" requires clarification for readers unfamiliar with object detection methods. A brief explanation has been added to the manuscript at this point.

Changes in manuscript: The following text has been revised in the manuscript at Line 164:

"This eliminates the need for predefined anchor boxes, which are fixed rectangular templates used as starting references for object localisation, that were used in previous YOLO versions and several other detection methods"

Comment 7: Line 107. Define InSAR.

Response: The definition of InSAR has been included in the revised manuscript as follows.

Changes in manuscript: The following text has been revised in the manuscript at Line 170:

"In their study, landslide detection was performed using InSAR (Interferometric Synthetic Aperture Radar, a technique that uses radar signals to measure ground surface deformation) measurements, where YOLOv8n base model achieved a detection performance of 96.76% mAP50."

Comment 8: Line 109. "detection accuracy of 96.76% mAP50". What does this mean?

Response: As defined in our response to Comment 5, mAP50 is the mean Average Precision evaluated at an Intersection over Union threshold of 0.50. Here, an mAP50 value of 96.76% indicates that the model correctly detected and localised 96.76% of the true landslide targets in the dataset at an IoU threshold of 0.50.

Changes in manuscript: No changes to the manuscript text are required for this comment, as the mAP50 definition has already been incorporated in response to Comment 5.

Comment 9: Line 113. What are "convolutional layers"?

Response: A brief explanation of convolutional layers has been added to the revised manuscript at this point.

Changes in manuscript: The following text has been revised in the manuscript at Line 181:

"It has a series of convolutional layers, which are mathematical operations that scan the image using small filters to detect local patterns such as edges, textures, and shapes."

Comment 10: Line 154, "The model was trained for 50 epochs". What is meant by an epoch?

Response: The term "epoch" is specific to machine learning and may not be familiar to readers from the atmospheric science community. In the context of neural network training, one epoch refers to a single complete pass of the entire training dataset through the model. A brief clarification has been added to the revised manuscript at this point.

Changes in manuscript: The following text has been revised in the manuscript at Line 220:

"The model was trained for 50 epochs, where one epoch represents a single complete pass of the training dataset through the model."

Comment 11: Table 1, what do "optimizer" and "learning rate" mean?

Response: We agree that these terms require clarification for readers outside the machine learning domain. Brief definitions of both parameters have been added to the table caption in the revised manuscript.

Changes in manuscript: The following text has been added to the caption of Table 1:

"The optimizer refers to the algorithm used to update the model parameters during training to minimise prediction errors. The learning rate defines the step size of these updates."

Comment 12: Line 158, define "confusion matrix"

Response: The confusion matrix has been defined in the revised manuscript at this point.

Changes in manuscript: The following text has been revised in the manuscript at Line 225:

"Precision and recall values are derived from a confusion matrix, which is a tabular summary of model predictions categorised into true positives (TP) that represent correctly detected LIME signatures, false positives (FP) indicating wrong detections, and false negatives (FN) which denote missed LIME signatures."

Comment 13: Line 209: "Training stabilised after approximately 30 epochs." What metric is used to determine this stabilisation. The slope does not appear to change from epoch 10 onwards so I'm not sure how stabilisation can be deduced?

Response: We acknowledge that the wording "stabilised after 30 epochs" could be misleading. The sentence has been removed and the manuscript text revised accordingly.

Changes in manuscript: The following revised text has been included at Line 269:

"Both losses exhibited rapid initial decrease during the first 10 epochs, followed by progressively smaller improvements in later epochs."

Comment 14: Line 227, "The detections above 70 km appear to be outliers possibly caused by interference." Is this the same kind of narrow-band interference that is seen at a Doppler shift of -0.35 in Figure 7. If so, it may be worth pointing this out so the reader can see what the interfering signals look like.

Response: This statement was indeed a bit superficial and we inspected detections near 70 km. Some of them were EPP-related echoes, others actually don't show the typical parabolic EPP image power spectra. The latter were often short-lived and near zero Doppler shift, making them difficult to interpret.

Changes in manuscript: We changed the sentence to:

"The detections above 70 km appear to be outliers, which can be faint EPP-related echoes without substantial absorption above, aircraft reflections or possibly also caused by interference."

Comment 15: Line 230: "representing the approximate boundary of daylight conditions". I suggest you add "at the altitudes of interest" to this line.

Response: This suggestion has been incorporated into the revised manuscript.

Changes in manuscript: The following text has been revised in the manuscript at Line 287:

"The red dashed contour marks the solar zenith angle of 95°, calculated at ground level at the radar location, used here as an approximate indicator of the daylight conditions at the altitudes of interest."

Comment 16: Line 230, "The vast majority of detections". You later quote the number of detections occurring doe SZAs above 95 deg as 99%, so I suggest you add "(99%)" after "detections" so the reader does not have to get further into the paper to find out this number.

Response: This is a valid suggestion. The value "(99%)" has been added after "detections" in the revised manuscript.

Changes in manuscript: The following text has been revised in the manuscript at Line 289:

"The vast majority of detections (99%) occurred when the solar zenith angle remained below 95°, indicating the necessity of solar illumination for the generation of LIME."

Comment 17: Line 255, "While the quasi-simultaneous precipitation of similar energies occurs, and thus corresponding altitudes". I don't understand what the authors mean by this. I strongly suggest this sentence is re-written to make it clearer.

Response: We've indeed been too quick on this.

Changes in manuscript: We rephrased and extended this statement within Lines 315 – 330.

Comment 18: Line 336, "prioritisation of detection reliability over completeness". What do the authors mean by "completeness". Do they mean visually checking all the spectra for LIME?

Response: By "completeness" we refer to recall, i.e. the proportion of true target signals (LIME signatures) that are correctly detected by the model. In this context, prioritising detection reliability over completeness means that the model favours reducing false positives (higher precision) at the expense of missing some true detections (lower recall). This does not refer to manual visual checking of spectra. The manuscript text has been revised to clarify this terminology.

Changes in manuscript: The following text has been revised in the manuscript at Line 415: *"The higher precision (89%) than recall (80.4%) indicates that the model favours reducing false positives at the expense of missing some true LIME detections"*

Suggestions for improved readability

Response: We thank the reviewer for the numerous suggestions to improve readability and language throughout the manuscript. All suggested wording and grammatical corrections have been carefully considered and incorporated into the revised manuscript.

EGUSPHERE-2026-1030

Automated Detection of Low-altitude Isolated Mesospheric Radar Echoes Using YOLOv8: Evidence for a C-Layer Phenomenon near 60 km Altitude?

Yadu Krishnan Krishnakumar, Toralf Renkwitz, and Andreas Ahrens

Answers to Reviewer #2

We thank the reviewer for the careful reading of the manuscript and the valuable comments provided. Please find our responses to the comments below. We're confident to have improved the manuscript on the raised issues.

The authors should remove any claim of a "C-layer phenomenon" from the title and body of the manuscript. The presentation of the work should be reoriented towards the important central findings (occurrence rates, characteristics etc.) and away from process issues (choice of computer tool etc.) and speculative conclusions.

Response: We thank the reviewer for this statement. However, for both raised points the authors believe that they are important and should be stressed. To better explain and motivate this, we've rewritten substantial parts of the manuscript. Our explanation indeed was speculative, but we aimed to exclude two other possible processes, that could cause the radar echoes at the lower D-region. The machine learning approach is employed and described to reliably detect the often weak radar echoes following an initial manual and thus time-consuming search.

Changes in manuscript: The manuscript has been widely modified, especially in the Introduction, but also the Discussion and Conclusions. We also added more convincing material as GCR detection rates.

Major comments:

Comment 1: Novelty claim is overstated

Previous studies cited by the authors have reported similar low-altitude reflecting layers (e.g., Rasmussen et al., 1980; Bain & Kossey, 1987). The authors should reframe the novelty as systematic detection and characterization rather than first discovery – in particular the term "first evidence" should not be used.

Response: This criticism relates to our statement in the abstract: "To our knowledge, this study might be the first evidence of such a phenomenon through consistent radar observations." This was indeed too vague and imprecise.

Changes in manuscript: We have rephrased this specific sentence to: "To our knowledge, this study might be the first systematic investigation of such a phenomenon through monostatic radar observations." We also substantially extended the Introduction, modified and extended the Discussion section to clearly differentiate between the multistatic (V)LF observations inferring a "C layer" and the monostatic, vertical sounding and range unambiguous observations we report here.

Comment 2: Physical interpretation is speculative

The proposed link to galactic cosmic rays is not sufficiently supported. Correlations with solar flux are weak and no quantitative ionization modeling is presented.

Response: Indeed, the correlation of solar flux is weak, basically only little correlation for solar minimum years, but none for solar maximum. We also agree, that the link to galactic cosmic rays was speculative. Modelling of the ionization scenario is far beyond the scope of this report, but we'll aim to initiate such an effort.

Changes in manuscript: The correlations for solar flux and LIME are given in Line 342. We have now added GCR detection/occurrence rates to Fig. 12 (in the revised version). For 2024, the reduction of GCR during solar active periods is very prominent including quasi-simultaneous absence of LIME. We also adapted the discussion accordingly.

Comment 3: Limited training dataset

The YOLO model is trained on only 200 images, which may introduce selection bias and limit generalization.

Response: We acknowledge that 200 annotated images represent a relatively small training dataset. However, as described in the manuscript, we employed a transfer learning approach using a pre-trained YOLOv8n model, which already provides robust feature extraction capabilities learned from large-scale datasets. Fine-tuning such models on smaller, domain-specific datasets is a well-established practice. Nevertheless, we agree that increasing the size and diversity of the training dataset would further improve the model's generalisation performance in future work.

Changes in manuscript: No changes were made to the manuscript for this comment, as the methodology and discussion already address the use of transfer learning and its applicability to small datasets.

Comment 4: Subjective ground truth

Manual labeling introduces subjectivity. Clearer quantitative criteria for LIME identification are needed.

Response: We understand that the manual labeling process is subjective and that explicit quantitative criteria for LIME identification were not formally defined in this study. The annotation was guided by the primary visual characteristic of LIME, isolated radar echoes appearing at approximately 60 km altitude with clear spatial separation from D-region echoes at higher altitudes. As this work represents a first application of automated object detection to LIME identification, establishing a fully objective, quantitative annotation protocol was beyond the scope of the current study.

Changes in manuscript: No changes were made to the manuscript for this comment, as the limitation is already inherent in the described annotation approach.

Comment 5: Single-instrument limitation

All results rely on one radar system, limiting generalizability. How do these results compare to VLF/LF data, or to data from optical/other instruments? A 3 MHz plasma layer (equivalent to 10^5 el. m-3) should be easily visible using ionosonde data, with suitable processing. Particle or FUV data (see e.g. DMSP) could be used to test for energetic precipitation.

Response: This is indeed correct, and somewhat regrettable. The only other clear example we could find is shown in Reid 2015. Besides this example, we could not find any reports on LIME for other MF/HF PR radars, nor could we identify them in publications. The much lesser power-aperture product of the compact radars might be the major limitation. Also, other instruments like Ionosondes will suffer the same limitation and the typically start receiving at 80 km at earliest as very long codes are used to compensate for the compact antenna setup.

Changes in manuscript: Besides the extension of the Introduction, we also we address this point in Line 428.

Comment 6: Alternative explanations

Other mechanisms (e.g., turbulence, gravity waves) are not sufficiently explored. How do we know this scattering is really caused by ionization, and that the ionization is unrelated to energetic particle precipitation? Could it be mono-energetic precipitation?

Response: We certainly do not rule out any contribution of dynamics, they may actually be important for the species that are ionized. This, however, is far out of the scope of this manuscript, but we aim for a subsequent manuscript proposing a model-based explanation. The partial reflection radar echoes sufficient gradients of electrons density per unit volume are required. Turbulence and also atmospheric gravity and other waves (tidal etc.) are certainly present. The observed spectral detection box width enhancements during October might indeed be an indication for enhanced turbulence and if that's reliable will open a new valuable dataset by itself.

We’re not certain if mono-energetic precipitation events may occur. However, assuming they can, corresponding energies (500 keV and above) would cause significant absorption reaching and partially reflecting at the D region, which should be visible. In order to clearly separate between LIME and EPP-related radar echoes, we remove periods of detected EPP.

Changes in manuscript: We have added a statement in the Conclusions as an outlook for a follow-up study.

Minor comments:

Comment 7: Improve figure labeling (units, descriptions etc), clarify uncertainties, check spelling and grammar.

Response: We thank the reviewer for the helpful suggestions to improve clarity and presentation. All figures have been checked and updated for consistent labeling, including units and clearer descriptions. Uncertainty statements have been clarified where needed, and the manuscript has been carefully proofread to correct spelling and grammatical errors.

Changes in manuscript*:* Relevant updates have been made throughout the manuscript.

Comment 8: why was a ~63-km layer chosen for Fig 1a when the abstract claims most of these are ~58 km? The claim in the abstract should be rephrased in terms of mean and standard deviation of the LIME detection altitude – Fig 9 makes it clear that a relatively broad range of altitudes is present.

Response: This example has been chosen as it is one of the strongest events and with the longest duration. It also nicely depicted the separation to the regular D region echoes.

Changes in manuscript: We are now presenting another example in Fig.1, which more closely follows the statistics. We also updated the abstract to contain more findings of this study.

Comment 9: What do the authors make of the apparent October-March (and then June-August) concentration of detections, in terms of physical mechanisms? Doesn’t it look like two separate phenomena, considering also the spectral width variation? If so, maybe two terms are needed rather than just “LIME” for both?

Response: There are indeed substantial seasonal differences in altitude and spectral widths. The general appearance, however, is fairly unchanged and also the requested isolation to the D region. We are hesitant to introduce two different classes as of now as it might be rather the seasonal imprint of existing turbulence.

Changes in manuscript: The seasonal differences of spectral widths are discussed around Line 300.

EGUSPHERE-2026-1030

Automated Detection of Low-altitude Isolated Mesospheric Radar Echoes Using YOLOv8: Evidence for a C-Layer Phenomenon near 60 km Altitude?

Yadu Krishnan Krishnakumar, Toralf Renkwitz, and Andreas Ahrens

Answers to Reviewer #3

We like to also thank reviewer #3 for the careful reading and suggesting valuable improvements for the manuscript. Some of the raised points are in alignment with previous comments of the other two reviewers, which is good to strengthen those topics explicitly. Our responses to the comments are provided below.

… The description in the manuscript is, however, largely imbalanced, with too much emphasis on the machine learning technique and at the same time lacking sufficient citations to the existing references, technical background of the radar system and also quantitative science discussion. I therefore recommend a major revision before the manuscript becomes suitable for publication.

We realize the amount of Machine Learning (ML) content is still rather unexpected for this journal, however, the application of ML is listed in the addressed topics. Also corresponding to the previous reviewer comments, we want to stress the potential and performance of this ML approach in comparison of very time-consuming manual inspection and labelling of many thousands of spectra.

We of course agree on improving the manuscript, adding more references, refining the introduction and clearly separating the individual sorts of mesospheric radar echoes that may be observed, including a differentiation to the here reported LIME. We'll address the individual points subsequently.

Major comments:

Comment 1: A significant part of the manuscript is for the explanation on how to use YOLO and its usefulness in the 'LIME' detection. Because EGUsphere is a journal of Earth, Space, and Planetary sciences, the description for the machine learning technique should be more compact and efficiently summarized, and more science discussion should be made instead. That said, I understand that such machine learning techniques are very useful in geoscience studies. If the description of technical details of the machine learning approach is a major point of this work, the paper should be separated into two, one for a technical report and the other one for a science paper.

Response: We recognise the imbalance between the machine learning description and the scientific discussion in the original manuscript. The YOLOv8 methodology section has been revised to be more compact by reducing technical details while retaining the information necessary for reproducibility and clarity. At the same time, we note that a certain level of methodological description is important, as this represents a relatively new approach for LIME identification. We believe the revised manuscript now provides a better balance between methodological explanation and scientific discussion. We don't intend to split up the content into two publications, however, we plan for a follow-up paper including more data and analysis also from other radar experiments.

Changes in manuscript: The methodology section describing the YOLOv8 framework has been shortened and streamlined to reduce excessive technical detail. Additional emphasis has been placed on the scientific interpretation and discussion of the LIME observations throughout the widely revised manuscript.

Comment 2: In the YOLO detection the definition and the threshold of the following values are quite vague probably due to the nature of machine learning approach.

- definition and threshold of LIME, VOID and D region echoes
- height width
- frequency width
- height separation between LIME and D region echoes

Response: We acknowledge that explicit thresholds for the definition of LIME, Void, and D-region echoes as well as for height width, frequency width and height separation between LIME and D-region echoes were not formally defined in this study. This is inherent to the nature of the machine learning employed. During manual annotation, LIME signatures were identified as isolated radar echoes appearing at approximately 60 km altitude with sufficient spatial separation from D-region echoes at higher altitudes. To ensure this separation, the area with no signal presence directly above LIME was annotated as the "Void" class while the LIME signatures themselves were labeled as "Target_Signal" (see Fig. 4). Rather than relying on manually defined thresholds, YOLOv8 learns to identify and localise these features directly from the annotated training data. During training, the model is repeatedly exposed to these labelled examples and adjusts its internal parameters to recognise the spatial and spectral patterns associated with LIME echoes. The detection boundaries and separations are therefore implicitly learned from the data rather than explicitly prescribed.

Changes in manuscript: No direct changes were made to the manuscript regarding this comment, as the current methodology section already describes the annotation procedure and the YOLOv8-based detection approach.

Comment 3: Once the detection is made, I would re-evaluate these widths and separation by a fitting technique with a clear threshold setting to avoid vagueness. What the authors claim 'frequency width' is the apparent region where the frequency enhancement is seen, not the same with the width related to turbulence activity. It is also not known if the spectral width of MF echoes is a simple measure of turbulence activity. The frequency widening due to turbulence is usually evaluated by a Gaussian fitting to the MST radar spectrum, values evaluated through which are independent of echo intensity. The description around Line 276 is therefore inappropriate although turbulence can affect MF spectra.

Response: That's correct, the given widths are not necessarily the widths of the echoes, but rather of the detection box (Lines 277, 296) and is therefore prone to an overestimation. Nevertheless, an increase in detection box width will relate to spectral width. How this MF/HF spectral width then relates to turbulence still needs to be verified. One part of the planned follow-up paper is to use the validated detections and derive the spectral widths by e.g. gaussian fitting. The derived spectral widths, however, need to be compared to the normal (solar/geomagnetic quiet) D region MF/PRR widths at the same time. Such a statistic for Saura has not been published so far, but is planned for the near future. Addressing both in this manuscript will certainly overload this paper.

Changes in manuscript: Throughout the manuscript, we added notes, that the shown widths are from the detection bounding boxes.

Comment 4: In Introduction the authors need to explain more about the history of low altitude mesosphere echo studies in the polar region, especially in winter, adequately referencing previous studies both in the Arctic and Antarctic such as Hall et al. , ACP (2006), Morris et al., GRL (2011), Renkwitz and Latteck (2017), Nishiyama et al. , JGR (2018) and other PMWE related papers before they declare at around the line 38 that they deal with non-EPP type echoes. The authors need to make the differences even clearer between the existing works and the current work if there are such differences. Because the study of low altitude mesosphere echoes in the polar region has a rather long history, readers will be confused and may not be able to understand the most important point the authors claim.

Response: As we quoted earlier, also for the other reviewers, we realised and agrede on improving the introduction, introducing the typical and quiet D region and differentiate more clearly between the individual sorts of radar echoes and corresponding references. We've been clearly too short on that part with the initially submitted version.

Changes in manuscript: We have made substantial changes throughout the manuscript, specifically in the introduction, discussions and conclusions.

Comment 5: The term LIME introduced in the present study is awfully confusing and inappropriate. Hall et al. (2006) introduced ILME (Isolated Lower Mesosphere Echoes) for MF radar echoes observed under disturbed conditions over Tromsoe (69N), which were thought to be strongly related to VHF radar echoes widely known as PMWEs. LIME and ILME are composed of the initials of exactly the same 4 words with only a different order. The term ILME has been used in the radar studies since Hall et al. and appeared at least in several papers to my knowledge, including Renkwitz and Latteck (2017). According to what the authors claim, these two abbreviations correspond to different background conditions, that is, disturbed and quiet. If so, this confusing naming should be avoided. I also feel that introducing LIME without mentioning Hall et al. (2006) is hugely disrespectful to the late Prof Hall.

Response: The not mentioned reference to Hall et al. (2006) is indeed very unfavourable, but certainly wasn't meant to disregard this publication, nor the authors or the involved research groups.

Considering the proposed name of these radar echoes, we don't share the criticism though it's indeed fairly similar to the "ILME", that Hall et al. (2006) introduced. "LIME" is still quite different to "ILME" from the visual appearance as well as the spoken words. Generally, we are open to other suggestions, but we actually like the term as it accurately describes the phenomenon without adding too much speculations to the actual origin.

Changes in manuscript: We have included the Hall et al. (2006) reference, when this sort of radar echoes is described and discussed.

Comment 6: The technical details of the Saura MF system are missing, which are only briefly mentioned in Conclusions. It is one of the largest MF radars together with the Adelaide radar, and only large one in the polar region. The sensitivity is thought to be significantly higher than the other existing conventional broad beam systems, making much easier the detection of lower altitude echoes even under non-EPP conditions. The sharp beam operation is also advantageous in suppressing unwanted off-vertical echoes resulting in a better height resolution with much less range smearing. Such technical differences from the conventional systems should be emphasized quantitatively in Introduction to claim the uniqueness of the current study.

Response: We are not completely sure which technical details of the Saura radar are missing. Section 2 describes the radar setup and covers the prime parameters. We, however, have added some more details, e.g. the nominal beamwidth and experiment sequence. The nominal narrow beam of Saura is indeed advantageous, but as it's a sparse antenna array, it still has significant radiation sidelobes. These sidelobes do not pose much a problem for normal D region echoes, but they clearly do during EPP-events.

Changes in manuscript: We have added more information on the radar system in Section 2 - Instrumentation and data. We also sketched the differences to other compact MF/HF PR radars in the Discussion.

Comment 7: The conditions for MF echoes to be detected should be mentioned and discussed more clearly by citing appropriate manuscripts: enough electron density to increase the refractive index, something such as atmospheric turbulences to fluctuate the index, and possibly something to reduce the mobility of ions such as NOx for coherent detection. The difference of O and X modes also needs to be explained together with the reason of selecting X mode in this study (e.g., Renkwitz and Latteck, 2017; Vierinen et al., 2013). While such background information is only very briefly scattered in the present manuscript and mentioned some in Lines 316-321, it should be more collectively summarized, perhaps in Introduction. Additional information on how O mode echoes look like will also be helpful for a better understanding.

Response: We agree that we needed to introduce the regular D region radar echoes and involved processes and relevant references. The same has been done for EPP-related events. The extraordinary magneto-ionic mode (X-mode) has been chosen as it's typically less affected by interference from other transmitters. It also is the preferrable mode for the lower altitudes at this radar position. We could not see any substantial difference in the O-mode echoes to the X-mode ones so far, but we aim to investigate this more closely in a subsequent study.

Changes in manuscript: We have made substantial changes specifically in the introduction, discussions and conclusions including corresponding relevant references.

Comment 8a: Regarding the refractive index fluctuations, turbulence activity is a key as the authors briefly mention. Atmospheric waves are believed to be largely responsible for the turbulence generation. Because of the nature of atmospheric stability, turbulences are generated more preferably at unstable phases of atmospheric waves, not necessarily all through the existing height region of those waves. MST radar observations, usually with better height resolutions than MF radars, often show layered structures as seen in Figure 3 (d) of Nishiyama et al. (2018) (see the two layered structure on May 28, 2013). Keeping this in mind, are there any possibilities that the isolated structure of 'LIME' is, at least partly, related to this layered turbulence structure?

Response: For the given time series shown in Fig 3c of Nishiyama et al. (2018) we agree on the authors statement that these echoes are primarily EPP-related echoes, ILME as Prof. Hall called them. In these examples multiple consecutive days show large absorption of the MF radar signal above 60-70km altitude during daytime. The D region electron densities are substantially enhanced allowing ILME to be seen in the MF and most of the normal D region is absorbed. Due to this strongly increased electron densities PMWE echoes are then visible in the MST radar – Fig 3d. This scenario is exactly different to what we refer to with our detections called LIME, where the normal D region is still visible and often clearly separated to the LIME. This means, the electron densities are not elevated throughout the D region, but small enhancements and suitable gradients near 58 km are required. Still, similar waves may of course be present for both, that are visible in the PANSY PMWE, and possibly for LIME times.

Changes in manuscript: The reference Nishiyama et al. (2018) has been added, when the connection of EPP-related echoes and PMWE are presented. EPP

Comment 8b: I presume that under moderately disturbed conditions (weak EPP) an MF radar, especially the sensitive Saura radar, could measure echoes from a wide height region (50-90km) with stronger echoes at unstable heights, resulting in an apparent structure seen as 'LIME'. The isolated echoes around 55 km of MF radar on May 29, 2013 seen in Figure 3 (c)of Nishiyama et al. (2018) might be such an example, where the MST radar echoes are only weakly detected in Figure 3 (d).

Response: As mentioned just before, we would consider this echo as ILME due to EPP, allowing for enough electron densities to make the structures visible as PMWE in the MST radar.

Comment 9: In the VOID areas seen in Figures 1, 4 and 7 of the manuscript, I see weak echoing region between LIME and D region echoes. As I mentioned, this kind of patchy structures are quite common in sharp beam MST radar measurements. So it would also be a case for a powerful high resolution MF radar such as Saura system.

Response: We are not certain which patchy structures are referred to in our Fig. 1, 4 and 7. There are a couple of weak spectral lines, which partly spread over multiple if not all ranges. These are certainly artefacts and possibly a weak DC-offset at 0 Hz.

Changes in manuscript: As requested by Rev.2, we have exchanged the earlier Fig. 1 with a different LIME example with the time series power and spectra.

Comment 10: GCRs can be a steady source of the low altitude echoes. References about ionization will be necessary around Line 323. However, the discussion about GCRs is still only speculative. If the same kind of echoes are not detected at other sites, especially at Juliusruh, necessary conditions for MF radar detection may be more quantitatively estimated considering the technical (power/gain) and latitudinal (GCR) differences.

Response: We thank for this statement and as we also pointed out to Rev.2, other detections of the same type of echoes would be very advantageous. The only other example we found is shown in Reid (2015), from a comparably powerful MF PR radar in Buckland park, Australia. All other vertical sounding MF/HF PR radars might actually be not sensitive enough. We could not find any example for Juliusruh PR yet. Despite this radar has equivalent peak pulse power, the antenna array is significantly smaller, thus it has a wider beam. This wider beam is even broadened by amplitude tampering in order to damp the existing sidelobes.

Changes in manuscript: We added references for GCR and their ionisation relevance as well as included detected GCR in the now Fig. 12. We also added a discussion on the sensitivity of the radars.

Minor comments:

Comment 11: For grammatical errors or suggestion, see comments by other reviewers.

Response: We thank the reviewer for the suggestion. The grammatical corrections and language related improvements suggested by the other reviewers have been carefully considered and incorporated throughout the revised manuscript.

Changes in manuscript*:* The manuscript has been thoroughly proofread and revised to improve grammar, wording, and overall readability throughout.

Comment 12: Figure 4 - Is the grey hatch necessary?

Response: We appreciate the reviewer's observation. The grey hatch in Fig. 4 is part of the default visualization style generated by the Roboflow annotation interface and was retained in the figure. It helps visually distinguish the Void class from the Target_Signal class.

Changes in manuscript: No changes were made to the figure, as the grey hatch originates from the default Roboflow annotation visualisation and aids class differentiation.

Comment 13: Figure 8 - Height width distribution is also wanted together with the separation distribution (VOID width).

Response: We thank the reviewer for this suggestion. Due to the scope of the current revision, the height width and separation (Void width) distributions have not been included in the present manuscript. We agree that these additional distributions would provide further quantitative characterisation of the detected LIME signatures and will consider including them in future work.

Changes in manuscript: No changes were made to the manuscript regarding this comment. However, the suggestion has been noted for future analysis and extension of the study.

Comment 14: Figures 9 and 10 - Is the 95 degree measured on the ground? If so, the value will be significantly different in the mesosphere. A clear definition will be wanted to avoid confusion.

Response: Thank you for this important observation. The solar zenith angle (SZA) of 95° used in this study was calculated at ground level using the geographical coordinates of the Saura radar (69.1°N, 16.0°E) with the astropy Python package. We acknowledge that the effective SZA at 60 km altitude might be slightly different. Therefore, the 95° contour is intended only as an approximate indicator of the day/night transition rather than a strict physical threshold. This clarification has been added to the relevant figure captions.

Changes in manuscript: The following sentence has been added to the captions of Figs. 7, 8, and 9:

"The red dashed contour marks the solar zenith angle of 95°, calculated at ground level at the radar location, used here as an approximate indicator of the daylight conditions at the altitudes of interest."

Comment 15: Figure 11 - What is the maximum value? Detection rates may be more understandable rather than actual number. Does the median mean one of the 4 values corresponding to the 4 years? Why not average?

Response: We acknowledge the reviewer's questions regarding the interpretation of Fig. 11. The maximum median occurrence value shown in Figure 11 is 22 detections per 30 minute time bin. We acknowledge that detection rates could be an alternative representation; however, we opted to show absolute counts because each detection corresponds to a confirmed LIME event, making the number intuitively interpretable.

The median is calculated across the four years (2018, 2019, 2020, 2024) for each week–time bin, not selected from a single year. For example, if a bin had yearly occurrence counts of [5, 8, 12, 20] for the four years, the plotted value would be the median (10), not any one year's value. The median was chosen over the mean because it is robust against outliers, years with unusually high or low detection counts (e.g. solar maximum year 2024) would otherwise skew the mean.

Changes in manuscript: The caption of Fig. 9 has been revised to clarify that the plotted values represent the median occurrence calculated across the four observational years for each week–time bin.