Dear Referee,

We would like to thank you for taking the time to review our paper. We reply to your comments below. Our response to the comments appears in bold and revised text as *italic*.

- L37-39: Drawbacks of these methods are that they either have a relatively low temporal or spatial resolution and that these are indirectly relating surface characteristics to evaporation.
  → Drawbacks of these methods are that they have relatively low temporal or spatial resolution, and they indirectly relate surface characteristics to evaporation.
  We followed the suggestion of the reviewer. We only left out the added comma after resolution, as we think it should not be inserted here.

- L42-45: long sentence that could be cut in two to easy the reading
  We followed the suggestion of the reviewer. The new sentences are:
  As a consequence of the different temperatures and humidities of turbulent eddies, density varies spatially and temporally and thus also the refractive index. This causes the signal intensity at the receiving end of the propagation path to fluctuate in time (typically at time scales between 0.1 and 100 s).

- L56-57: please rephrase to simplify
  We rephrased as follows:
  If we would be able to successfully use them as scintillometers, it would mean that we can estimate rainfall and evaporation with a single experimental setup, similar to Leijnse et al., (2007b, c).

- L57: remove "the"
  We followed the suggestion of the reviewer.

- L74: Add comma after "theory"
  We followed the suggestion of the reviewer.

- L76: "the Netherlands" between brackets
  We followed the suggestion of the reviewer.

- L97: Please precise this is for large aperture
  Tipically, large aperture (root_square(lamba . L) >> D) scintillometer 3D-power spectra ...
  We respectfully disagree with the reviewer here. The Kolmogorov 3D refractive index spectrum is assumed to hold for all apertures. Therefore, we did not follow the suggestion of the reviewer. See for example:
  Beyrich, F., Hartogensis, O. K., de Bruin, H. A. R., and Ward, H. C.: Scintillometers, in: Springer Handbook of Atmospheric Measurements, edited by Foken, T., pp. 969–997, Springer International Publishing, ISBN 978-3-030-52171-4, https://doi.org/10.1007/978-3-030-52171- 4_34, 2021

- L98-102: please cut this sentence in two.
  We followed the suggestion of the reviewer:
  For a power spectrum of intensity measurements obtained from a scintillometer with a given setup, the power spectrum depends on C_{nn} and $u_{\perp}$. Higher C_{nn} values increase the spectral

density over the entire range of scintillation frequencies, while higher $u_\perp$ values shift the scintillation spectrum to higher frequencies, while retaining the variance (e.g., Medeiros Filho et al., 1983; van Dinther, 2015).

- L104: please precise
  For small aperture scintillometers
  Or
  Neglecting Aperture terms in (1) and integrating this equation ...
  We again disagree with the reviewer. As stated in the paragraph after Eq. (2), the constant *c* depends on the experimental setup through the influence of the aperture averaging effect. Also the terms large-aperture versus small aperture, typically used for optical scintillometers, is ambiguous at microwave wavelengths. For optical wavelengths, large aperture scintillometers were introduced as saturation resistant and inner-scale independent type of scintillometers (Wang et al, 1977) by adopting a finite aperture over which small scale scintillations were averaged out. The consequence was that D should be used as the relevant eddy-scale instead of F as D>>F. Microwave scintillometers do have a finite aperture, yet F>D, so they can be considered small aperture scintillometers from that perspective. However, 1) they are saturation resistant, 2) inner-scale independent and 3) the aperture averaging effect is not negligible for microwave scintillometers which have a relatively short path and/or wavelength or a large antenna diameter, i.e. *c* in Eq. (2) is not a constant. For our microwave scintillometer, *c* is 2.60 instead of 2.01 if the path averaging were to be ignored. To conclude, we chose not to refer to our microwave instruments as small or large-aperture scintillometers in these paragraphs.

- L112-120: Good ! but I'm thinking the required precision in preceeding comments are still necessary to make it clearer.
  See our replies to your previous comments.

- L123: Remove "applying a"
  We followed the suggestion of the reviewer.

- L137: remove "the"
  We followed the suggestion of the reviewer.

- L207: "of the" ?
  We are not sure what the reviewer means here. We have slightly rephrased and added as follows:
  ...To do so, the power spectrum is smoothed by averaging the power at each frequency with those at the neighboring frequencies in a specified window. We specify the width of the window as 20 % of the specific frequency (to account for the increase in number of values towards higher frequencies in the power spectra). The weighting...

- L232: already precise above
  We agree with the reviewer and have removed this part.

- L246: remove "the"
  We followed the suggestion of the reviewer.

- L252: when ?

  We are not sure what the reviewer is referring to here. The RMBE is the mean bias between the logarithmic values on the y-axis and on the x-axis. By taking the logarithmic values, the RMBE can thus be interpreted as the difference expressed in terms of order of magnitudes. For example, an RMBE of 1 means that on average the values on the y-axis are 10 times greater than the values on the x-axis. For clarification, we have added:

  Intuitively, the RMBE represents the order of magnitude the values on the y-axis are larger (or smaller) than the reference values on the x-axis, due to the use of logarithmic values.

- L383: Personal comment : So much efforts on method 2 for very few values compared to method 1 !! This means that the main noise source is clearly between 1 and 10Hz.

  We think that the reviewer refers here to the difference in statistical metrics between the two, as the difference in number of observations is just 1.

  We agree that the improvement of the spectral noise method is relatively minor and does not reflect the extra amount of work required for this. Moreover, we are also aware that this method cannot be used in entire CML networks. However, we show this method as the "best possible" method to illustrate the potential, as we also argue in Section 6. Therefore, we leave this as is.

- L386: Also visible on Fig. 8:

  We added as follows:

  The RMBE related to both the MWS and EC has reduced from 1.2 to at least 0.3, which is major improvement in comparison to the RMBE of the comparison between the reference instruments (i.e., 0.01), indicating that the proposed methods overestimate C_{nn} at most with a factor 2 (i.e., $10^{0.3}$), which is also visible in Fig. 8a, where both methods are seen to overestimate the references during the entire day.

- L404: You can't say that. Metrics show very similar results and on cloudy day the spectral-noise method is much much noisy than the constant-noise method. I would accept "slightly better" but definitely not "outperforms" !

  We agree with the reviewer and have changed accordingly:

  …though the spectral-noise method performs slightly better than the constant-noise method,…

- L406: Remove "predetermining those parts of the spectrum behaving similarly to the" and add "the use of a"

  We agree with the reviewer and changed accordingly:

  …An advantage of the constant-noise method is that it is a relatively simple correction method which does not require the use of an MWS…

- L406-408: I'm still not convinced. Is it the frequency correction range or the wind range that imrpoves a bit the results compared to method 1 ?

  We refer here to the fact that, for the spectral-noise method, time intervals are distributed into different classes based on their crosswind, which in turn determines the frequency range over which the variance is computed. To clarify, we added as follows:

  Overall, this shows that considering crosswind conditions, which cause a shift of the power spectrum along the frequency axis, and selecting the frequency ranges over which scintillations are best resolved, also improves the C_{nn} estimation.

- L410: “which” → “this”
  We changed the order of this sentence:
  As the spectral-noise method requires the presence of an MWS to determine the contribution of noise to C_{nn}, the ability to transfer our methods to other datasets is limited.

- L423-425: Personnal comment : Note that accessibility to high frequency data will be necessary and imply discussions with CMLs operators. noise caracterisation for each CML version can then be part of a collaboration at least for intrinsic instrument noise sources (I mean not site sources),which are not difficult to measure in lab. Moreover operators could find an interest in suppressing noise from their network !
  We agree with the reviewer and have added an additional paragraph at the end of our conclusions as outlook:
  If the aforementioned challenges were to be successfully tackled, in theory it would be possible to estimate turbulent heat fluxes on close-to-continental scales, due to the large number of CMLs around the world. However, this would also require willingness from network operators and antenna manufacturers to support obtaining such spatial turbulent fluxes estimates. In the first place, it would be required to obtain signal intensity data from the network operators. Moreover, the currently most common sampling strategy in the network management systems is a minimum and maximum value of the received signal intensity once every 15 minutes. For turbulent flux estimates, it would be beneficial if the signal intensities would be stored at a higher sampling frequency, in order to be able to adequately estimate the variance per time interval. Lastly, close collaboration with CML antenna manufacturers would be needed to help understand and quantify the noise sources which are present in the different CML types.

- L431: Remove “similarly”
  We removed “similarly”.

- L434-444: The argumentation sounds a bit out of the scope. For exemple Cnn estimates from Penman to then calculate latent heat flux is not so much relevant.
  Your relevant point is at the end of this paragraph in term of spatial distribution of CMLs. I'm thinking you should shorten the first part and directly jump to the opportunity of the avalability of a large network to estimate CNN despite is uncertainties at the moment.
  Then the next perspective could be to evaluate the ability of such a network to map Cnn variability ...
  We realise that the aim of this paragraph was unclear. We aim here to compare our C_{nn} estimates with other methods. Therefore, we have restructured this paragraph by moving the last sentences to the front:
  Even though other studies outperform our C_{nn} estimates, these all require high-quality meteorological input data, which are not often available, whereas C_{nn} estimates obtained from CML signal intensities would be a more direct method to obtain C_{nn}, do not require any additional measurements and are available from a potentially larger number of devices with a nearly continental coverage. Van de Boer et al. (2014) used single-level observations to obtain the energy balance and used the Penman-Monteith equation to estimate C_{nn}. A comparison of their simulated C_{nn} estimates with EC-based C_{nn} estimates over grassland seems to outperform our comparison between CML and EC estimates, though their method shows a large dependence on the quality of the meteorological input data. Similarly, Tunick (2003) estimated C_{nn} using two-level meteorological observations of wind speed, temperature and humidity. Also, Andreas (1988) provided C_{nn} estimates over snow and ice by using meteorological

observations and emphasized the strong dependence of his estimates on the non-linear relation between the fluxes and C_{nn} and the dependence on the assumed Bowen ratio.

(Note that we moved our final sentence to Section 6.2, where it fits better than with the new structure of this paragraph, see next comment)

- L454-460: Other option could be to characterize component (antena, amplifier, detector, ...) noise source of a CML device
  The on you show on fig. 5 doesn't seem so difficult to model with continuous functions.
  We agree with the reviewer and have added as follows (also see our reply to the previous comment):
  ...having full information on the noise. Alternatively, the noise sources for each component of a CML device could be characterized in case high-frequency data is available. Moreover, the mounting mechanism of the CMLs is not designed to be vibration free, as the Nokia CML started to vibrate above 8 m s^{-1}, even though the mast itself remained free of vibrations. In addition, the used masts in CML networks might also not be vibration-free. Note that for long paths, saturation of the scintillation signal could also influence obtained C_{nn} estimates (e.g., see Meijninger et al., 2006, for the saturation limit for microwave frequencies).

- L499-505: This should be reported in the discussion section ?
  We are not sure what the reviewer refers to here. We have discussed this more elaborately in Section 6.2 and summarise this here, therefore we leave this as is.

- L511-513: May be to express this better. I don't know if this is what you mean but even though ML or else could be a way to work with high frequency scintillation data, they will require a huge amount of data for learning, and it will not suppress noise sources issue, potentially different from a CML device to the other.
  We agree with the reviewer and have added as follows:
  Lastly, an attempt could be made to directly retrieve information on the turbulent heat fluxes from the received signal intensities without following the scintillation theory. For example statistical methods or machine learning could be used for this, even though they will require a large amount of data, they might not be able to suppress noise in the received signal intensities, and they would probably differ per CML type.