We thank Reviewer 1 for very useful comments and suggestions which have significantly improved our manuscript. Our detailed replies are mark in bold below. Changes in the revised manuscript are highlighted in yellow.

Review to the manuscript
"A novel infrared imager for studies of hydroxyl and oxygen nightglow emissions in the mesopause above northern Scandinavia"
by P. Dalin et al.

The authors present a new imaging instrument that measures emissions from hydroxyl and molecular oxygen layers in the mesopause region. They describe the technical characteristics of the instrument as well as the derivation of temperatures from these measurements. The new instrument combines the measurements of IR emissions from two different molecules, because the centers of the two emission layers are located at slightly different altitudes. This allows tracing disturbances in the vertical direction additional to the horizontal domain which is enabled due to the imager technique.

Furthermore, they present the first measurements during the winter 2022/23. These measurements were compared with lidar and satellite observations in order to validate the temperatures derived from the imager measurements. Finally, the authors present some small case studies to illustrate the capability of the instrument to monitor temperature changes with time and to detect wave disturbances in both directions vertical and horizontal.
Generally, the manuscript is well structured and written and it addresses scientific questions within the scope of AMT. Thus, I recommend its publication after some minor issues are addressed.

1 General comments:
1. The imager shall be able to trace wave disturbances in the vertical domain. As the two emission layers (OH and O2) are not located at a constant altitude and the center altitudes vary with time and season, the distance between the two layers is also not constant. Thus, it should be difficult to obtain absolute information on the vertical propagation. Can you comment and discuss this possible limitation of the technique in some more detail.

These two emission layers (OH and O2) are indeed varied in space and time, making different height distances between these layers. At the same time, if the same wave package, having the same horizontal wavelength and observed phase velocity as well as propagation direction, is observed both in the OH and O2 layers, one can assume that the same gravity wave was propagating both in horizontal and vertical domains. According to the general theory of gravity waves (e.g., Gossard and Hook, 1975) a gravity wave propagates at some angle to the vertical, with tilted phase lines. This should result in an observed phase shift of the same gravity wave between the OH and O2 layers. Once a phase shift and horizontal wavelength are estimated from the OH and O2 maps, one can calculate the vertical wavelength by using the following relation:

$$\lambda_z=\lambda_x / \tan(\alpha)$$

where λ_z and λ_x are the vertical and horizontal wavelengths, α is the angle between wave phase lines and the vertical. Furthermore, if the buoyancy frequency N is a known quantity or is estimated by using lidar or satellite temperature profiles, one can deduce the intrinsic frequency ω of a gravity wave from the following relation:

$$\omega=\pm N\cdot\cos(\alpha)$$

Substituting known values of ω, N and λ_x into the dispersion relation for gravity waves one can estimate a vertical wavelength again, thus verifying the first estimation of a vertical wavelength. This method is valid for a limited number of gravity waves having vertical wavelengths less that the height distance between the two layers (about 7 km).

Another simple method of the estimation of a vertical wavelength of a gravity wave is based on the assumption that the height difference D between the two layers is a known quantity (Fagundes et al., 1995; Schmidt et al., 2018). If a horizontal phase shift $\Delta\varphi$ of a considered wave package between the both layers is estimated then one can calculate the vertical wavelength λ_z using the following relation:

$$\lambda_z = D \cdot 2\pi / \Delta\varphi$$

We will use the both methods to estimate vertical wavelengths of gravity waves propagating through the OH and O2 layers.

We have provided this information in the Discussion of the revised manuscript (lines 444-471).

2. Typically, the contamination of the lines, especially the P1(4) line, by other emissions such as emissions by the OH(4-2) R-branch is corrected during the temperature estimation process (e.g. Schmidt et al., 2013; Pautet et al., 2014). How is this contamination corrected for your measurements?

This is a complicated comment. At present, we do not correct the P1(4) line emission due to emissions by the OH(4-2) R-branch. We should note the following.

We could not find any information of the temperature correction due to emissions by the OH(4-2) R-branch in Pautet et al. (2014). Schmidt et al. (2013) just refer to the method proposed by Lange (1982) when discussing the temperature correction. Schmidt (2016), the doctoral thesis, compares Lange's correction with his own temperature corrections shown in Fig.2.15a in the doctoral thesis. The exact procedure of calculating these temperature correction factors has not been published. A discussion with Dr. Carsten Schmidt (personal communication) suggests that Lange (1982) did not provide details concerning the calculation of the temperature correction and this information is now lost. Temperature corrections by Lange's function as well as presented in the doctoral thesis by Schmidt (2016) were made for the R1(6) line using the value at the line center only. The OH imager registers an integrated intensity of the P1(4) line over a broader range with the interference filter having the spectral width (FMHW) of 2.1 nm. It means that the relative contribution by the R1(6) center line is higher compared to the total area of the R1(6) and P1(4) lines. Summarizing, Dr. Carsten Schmidt suggests that the contribution of R1(6) to the total area of the P1(4) line is less important and the temperature correction might be unnecessary.

Thus, at present, we could not find an explicit published procedure to correct our temperature estimations by the OH(4-2) R-branch for the case of the intensity integrated over the entire spectral width of the P1(4) line.

Also, we should note that a small temperature correction is not important for data analysis dealing with studies of wave disturbances and temperature seasonal changes. A small temperature correction might be important for the temperature validation with other instruments. But in this case, other instruments should also include the same temperature correction that is not obvious.

Finally, we should note that OH(3-1) rotational temperature estimations agree well with those by the Esrange lidar and Aura/MLS measurements as presented in the manuscript. We might correct the OH(3-1) rotational temperature in the future if we find a correction method appropriate for the OH imager.

3. Could you please clarify the name assignment of the different measurements, because it is a little bit confusing to me. In section 3.4 I_{P12} is introduced as the intensity of the P1(2) line, later it is called the raw intensity and in Fig. 4 this raw intensity is the sum of P1(2) and BG (background). Is I_{P12} the intensity observed in the spectral range of the the P1(2) line and P1(2) (in Fig. 4 and below Eq. 1) is the real line intensity of the line in counts after subtracting the background? And is the dark noise already subtracted from the measurements?
Maybe it is helpful to revise these names in the manuscript to get a clear and consistent name assignment.

We have revised the name I_{P12} and I_{P14} lines as being raw intensities, meaning that these raw intensities are without any correction, that is, no subtractions of the atmospheric background and noise have been made. We have provided this information in the revised manuscript (lines 238-240) as well as in the capture to Fig.2 (line 735).

2 Specific comments:
1. Eq. 2: Intuitively, I would expect that the background is subtracted from the measurements in each of the spectral ranges of the observed lines separately. Before this subtraction both single measurements (background and intensity in the spectral range of the emission line) should be corrected for the dark noise influence. Here the dark noise is added (with some factor). Can you please explain a little bit more where the equation comes from.

Equation 2 (new Equation 3 in the revised manuscript) comes from Equation 1 in the following way. The *R* brightness ratio $B(P_1(2))/B(P_1(4))$ in Equation 1 is in the absolute units (Rayleigh). The instrument registers emission intensities (P1(2), P1(4) and atmospheric background) in relative digital units (counts). In order to relate relative to absolute units, the absolute calibration is performed. The main part of this procedure is to determine filter absolute sensitivities which are different for each filter. These are the coefficients $k2$, $k3$ and $k6$ in Eq.3. In addition, the dark noise is subtracted both from the P1(2) and P1(4) lines as well as from the atmospheric background line which, in turn, is finally subtracted from the P1(2) and P1(4) lines. Since the coefficients k_2, k_3 and k_6 are different this procedure results in different constants (0.22 and 0.60) for the subtracted dark noise in the numerator and denominator in Equation 3. These constants have positive signs since the dark noise is subtracted from the $P_1(2)$ and $P_1(4)$ lines as well as from the atmospheric background line having different coefficients. The coefficients k_4 and k_7 describe the flat field correction factors being different for each OH (3-1) emission line. Finally, the coefficients k_1 and k_5 convert photometric units to the Rayleigh, which includes several multiplies and the geometric etendue.
We have provided this information is the revised manuscript (lines 246-258).

2. Eq. 3 and Eq. 4: Do the different coefficients k_i have their own uncertainties which then should be taken into account during the error estimation or are the uncertainties too small to have an impact on the total error?

Reviewer 1 is right, the different coefficients k_i have their own uncertainties which should be taken into account during the error analysis. We have provided new error estimations in the revised manuscript (lines 286-299 and new Equation 5).

3. Fig. 4: Maybe it is useful to also show the OH equivalent temperature which has been calculated from the lidar observation by vertical averaging in the figure.

We have added the average height-weighted lidar temperature (191.8±12.3 K) as was calculated by vertical averaging across the OH layer, shown by the green asterisk in Fig. 4 (lines 747-749 in the revised manuscript).

4. Fig. 6: It could be helpful to change the ranges of the colour bars as in most cases the full range is not present in the observations and some colours are not used then. This would maybe increase the contrast and visibility of the disturbances.

We have changed the ranges of the colour bars in Fig.6 in the revised manuscript.