Meteosat Third Generation Infrared Sounder: Sensitivity to measurement noise and spectral resolution in a non-linear variational retrieval framework

E. Pavelin, S. English, J. Eyre
Met Office, U.K.

Revision 0.1 23/1/06 EP
Revision 0.2 30/1/06 EP After comments from SE and JE
Revision 0.3 6/2/06 EP Added conclusions
Revision 0.4 8/2/06 EP After comments from SE
Revision 1.0 10/2/06 EP Final draft after comments from JE

1. Introduction

The aim of this study is to assess the impact of non-linearity on humidity and temperature retrievals from a proposed infrared sounder (IRS) aboard the future Meteosat Third Generation (MTG) geostationary satellite platform. Various scenarios are tested with differing instrument noise and spectral resolution characteristics in order to help determine the optimum design specification.

Measurements from the proposed sounder are simulated with a fast radiative transfer model (RTTOV), using a set of model profiles of atmospheric variables. Non-linear retrievals are carried out on the brightness temperatures using a stand-alone optimal estimation (1D-Var) package. The performance of the simulated instrument in various scenarios is assessed in terms of the mean errors in the 1D-Var retrievals compared to the original "truth" model profile dataset.

2. Study framework

2.1. Model profile dataset

In order to simulate a wide range of realistic atmospheric states, a large set of model profiles was used as the basis for the simulations of brightness temperatures and NWP background profiles. A set of 13495 model profiles derived from the ECMWF ERA-40 reanalysis (Chevallier, 2001) was chosen as the basis of this study. The profiles in this dataset were selected to be representative of a diverse range of atmospheric conditions, seasons, and geographical locations. The dataset contains profiles of temperature, humidity and ozone on 60 levels, as well as surface pressure, surface temperature, skin temperature, surface humidity, and surface type.

Figure 1, reproduced from Chevallier (2001), shows the variability in temperature and humidity profiles for the whole data set. For the purposes of this study, the dataset was interpolated onto 43 fixed pressure levels, corresponding to the RTTOV radiative transfer model levels used in the 1D-Var code.

2.2. 1D-Var Retrieval code

Retrievals were carried out using the NWP SAF Met Office 1D-Var code (Collard, 2004). It is a fully non-linear optimal estimation scheme based on the theory outlined by Rodgers (2000). The forward model used in the retrieval was RTTOV 8.5 (Saunders and Brunel, 2004).

The retrievals were configured to use a Marquardt-Levenberg minimisation scheme, which has been found to be the most suitable for non-linear problems.

The retrieval state vector consists of temperature on 43 pressure levels, humidity on the lowest 26 levels, skin temperature, surface temperature and surface humidity. No cloud variables are present in the state vector, so all retrievals are carried out under the assumption of cloud-free conditions.

2.3. Channel selection

It is assumed for the purposes of this study that the IRS sounder channels will be similar to those of AIRS. However, retrievals are also run with reduced spectral resolution and restricted spectral domains. For the "baseline" case, the sensor channel wavenumbers and spectral response functions were assumed to be the same as those of AIRS. 1589 AIRS channels (nos. 212-1800; 710–1560 cm^{-1} / 6.4–14.1 µm) were chosen, encompassing the main water vapour and carbon dioxide absorption bands, but omitting the short-wave AIRS channels. The sampled IR spectrum is illustrated in Figure 2.

2.4. Simulation of observations and model background

Top-of-atmosphere infrared brightness temperatures were simulated using the RTTOV 8.5 fast radiative transfer model (Saunders and Brunel, 2004). RTTOV is designed to simulate nadir viewing passive infrared and microwave satellite radiometers. Brightness temperatures were simulated for cloud-free conditions over all surface types represented by RTTOV.

Random observation errors were added to the simulated brightness temperatures with variances consistent with the three assumed noise scenarios. The prescribed brightness temperature noise was assumed to be valid for a reference scene temperature ($T_{\rm ref}$) of 280 K. The error added to each sounding was modified to take account of the actual scene temperature ($T_{\rm B}$):

$$\delta T_{\rm B} = \frac{{\rm d}R(T_{\rm ref})/{\rm d}T_{\rm ref}}{{\rm d}R(T_{\rm B})/{\rm d}T_{\rm B}}\,\delta T_{\rm ref}\,,$$

where R represents radiance, $\delta T_{\rm ref}$ is the BT error at the reference scene temperature, and $\delta T_{\rm B}$ is the BT error at the actual scene temperature.

The observation errors are Gaussian, uncorrelated, and unbiased. The observation error covariance matrix (R) used in the retrievals was calculated from the simulated noisy dataset for each noise scenario, and reflects variations in scene temperature over the whole dataset.

NWP forecast profiles were simulated for use as background input to the retrievals. The background profiles were calculated from the ECMWF profile dataset, and interpolated onto the 43 levels used in the retrievals. Simulated forecast errors consistent with the Met Office forecast error covariance matrix (B) were added to the ECMWF profiles to generate realistic background profiles. Background estimates of temperature, humidity, ozone (not retrieved), surface temperature, surface humidity, skin temperature, and surface pressure were supplied to the 1D-Var scheme.

3. Sensitivity to measurement errors

3.1. Observation error

Three instrument noise scenarios were proposed for this sensitivity study (summarised in Table 1). The "best case" scenario (N1) specified brightness temperature noise varying from 0.5K in the CO_2 band to 0.2K in the H_2O band, changing in steps of 0.1K across four bands (A – D). The "worst case" scenario (N2) has double the noise of N1. A "compromise" case

(N3) has low noise (0.2K) in the H_2O band (D) and double noise in the CO_2 band (A – C). These noise scenarios are illustrated in Figure 3.

Table 1: *Instrument noise scenarios*

	Wavenumber	Wavelength	N1 (Best)	N2 (Worst)	N3 (Compromise)
Band A	710 – 715 cm^{-1}	14.1 – 14.0 μm	0.5 K	1.0 K	1.0 K
Band B	715 – 730 cm^{-1}	14.0 – 13.7 μm	0.4 K	0.8 K	0.8 K
Band C	730 – 780 cm^{-1}	13.7 – 12.8 μm	0.3 K	0.6 K	0.6 K
Band D	780 – 2100 cm^{-1}	12.8 – 4.7 μm	0.2 K	0.4 K	0.2 K

1D-Var retrievals were run for the three noise scenarios using the full set of model profiles. In practice, only 12924 of the 13495 profiles were passed for retrieval because a number of profiles were affected by interpolation artefacts. The retrievals were run with the assumption of no forward model error.

Figure 4 shows the mean error characteristics of the full set of retrievals for all three noise scenarios. The mean retrieval biases and standard deviations relative to the truth are shown together with the same quantities for the background profiles. In addition, the analysis errors expected from linear theory are shown for the lowest-noise case. Specific humidity errors are expressed in terms of $\delta\ln(q)$, such that $\delta q = q\ \delta(\ln(q))$, where q is in units of g/kg.

In particular we note that there is only a small degradation in the performance of the humidity retrievals when moving from the "best" (N1) to "compromise" (N3) noise scenarios. This reflects the fact that the noise in the water vapour band in unchanged between these two cases. Temperature retrievals are more strongly degraded (by ~ 0.04 K) due to the increased noise in the CO_2 band. A more substantial negative impact on performance is noted in the "worst" (N2) noise case. In this case both humidity and temperature retrievals are affected to a significant degree. The temperature retrieval standard deviation is increased by ~ 0.14 K in the upper troposphere. The error in ln(q) increases by up to 0.03. There is no significant difference in retrieval biases between the three noise scenarios. Indeed, none of the runs presented in this study showed any significant differences in retrieval bias, so biases will not be discussed any further.

3.2. Forward model error

The three runs described in the previous section implicitly assumed zero forward model error, since the same forward model was used to perform the retrievals as was used to simulate the observations.

In order to assess the impact of a realistic forward model error, one run was carried out with a forward model error of 0.2 K added both to the R-matrix and to the simulated observations. Adding the error to the observations is equivalent to introducing the error at the 1D-Var forward model stage (and eliminates the need for changes to the 1D-Var code). The run was carried out with best-case (N1) observation errors.

Figure 5 shows the result of including forward model error in the N1 run. The effect of introducing a 0.2 K forward model error is small for both humidity (< 0.01 ln(g/kg)) and temperature (<0.02 K) retrievals. This result appears to justify the decision to neglect the FME in the other 1D-Var runs. However, the FME does introduce a small but consistent increase in the error standard deviation, so it should be noted that the results of the other 1D-Var runs shown in this study are slightly optimistic.

4. Sensitivity to spectral resolution

Since the IRS instrument is expected to have lower spectral resolution than AIRS, retrievals were carried out both with full AIRS spectral resolution (R1) and with half AIRS resolution (R2), in order to test the effect on performance of reducing the sensor resolution.

Half-resolution brightness temperatures were simulated by averaging together each pair of adjacent AIRS channels before adding noise. The reduced resolution was implemented in the 1D-Var retrievals by averaging the brightness temperatures and Jacobians returned from each forward model call in the same manner. Half the total number of channels was used (794 channels).

The three noise scenarios were run both at full AIRS resolution and half AIRS resolution. These runs were carried out using a subset of 3262 profiles (every fourth profile of the 13495-profile set). This smaller subset of profiles was used to reduce run time, as it was found that the error statistics did not change significantly when more profiles were included.

Figure 6 – Figure 8 show the effect of running with half-AIRS resolution for each noise scenario. Temperature retrieval standard deviations above 800 hPa are increased by no more than 0.1 K in all cases. The increase in the errors is largest in the upper troposphere. Specific humidity standard deviations increase by approximately 0.025.

5. Data denial experiments

Three additional runs were carried out with different portions of the CO_2 band excluded from the retrievals, using 12924 model profiles. Bands A, B, and C were excluded in turn, as shown in Table 2. The runs were carried out for the best-case (N1) noise scenario, with no forward model error and full AIRS resolution.

Table 2: *CO_2 band data denial runs*

	Bands excluded	Wavenumbers excluded	Wavelengths excluded
Denial 1	Band A	710 – 715 cm^{-1}	13.99 – 14.08 μm
Denial 2	Bands A + B	710 – 730 cm^{-1}	13.70 – 14.08 μm
Denial 3	Bands A + B + C	710 – 780 cm^{-1}	12.82 – 14.08 μm

The results of the data denial runs are shown in Figure 9. The effect on the humidity retrieval of excluding the CO_2 band is minimal. This suggests that there is sufficient temperature information present in the remaining portion of the spectrum to allow a good specific humidity retrieval. The temperature analysis is significantly affected by the exclusion of the CO_2 band, as expected. The temperature retrieval error increases by ~ 0.09 K in the mid-troposphere in the worst-case scenario (Denial 3), and this may be significant for the relative humidity analysis.

6. Summary and recommendations

6.1. Measurement noise

The sensitivity of the non-linear retrievals to changes in measurement noise has been explored using the three noise scenarios summarised in Table 1. The three scenarios consisted of a baseline “best case” noise specification (N1), a case with doubled noise (N2), and a case with doubled noise in the CO_2 band only (N3).

The effect of doubling the noise across the entire spectrum (N2 scenario) is to significantly degrade the quality of the analysis of both temperature and humidity. The quality of the analysis relative to the background is affected particularly strongly in the upper troposphere.

Doubling the noise in the CO_2 band only (N3 scenario) leads to a significant increase in temperature retrieval errors, especially in the mid- and upper-troposphere. The specific humidity analysis, however, shows only a very small degradation in quality. However, the increased temperature errors will have a negative impact on the quality of the relative humidity analysis, which is the important quantity in an NWP context.

In the best-case noise scenario, the temperature error $\delta T \approx 0.6$ K. At a fixed specific humidity of 0.5 g/kg at 600 hPa and 255 K, this translates to a 5.1% error in relative humidity. In the worst case ($\delta T \sim 0.7$ K), this error increases to 5.9%

In the absence of cloud, non-linear effects are most significant in the lowest-noise case. As the measurement noise is increased, the difference between the linear and non-linear analyses becomes less significant. This demonstrates that the linear analysis overestimates the performance gain from decreased measurement noise. In the worst-case noise scenario the linear and non-linear errors are almost identical. The differences between the linear analysis errors and the non-linear errors are largest below 600 hPa and in the upper troposphere (150 – 300 hPa).

6.2. Spectral resolution

In order to test the performance of an instrument with lower spectral resolution than that of AIRS (~ 0.5 cm^{-1} resolution, ~ 0.25 cm^{-1} sampling), the three noise scenarios were run with half AIRS resolution (~ 1.0 cm^{-1}). The temperature and humidity errors were affected to a similar degree in all three noise scenarios. In some circumstances the degradation in the quality of the analysis was similar in magnitude to the effect of doubling the measurement noise.

6.3. Longwave cutoff

The effect of removing the CO_2 band channels altogether is illustrated by the data denial runs plotted in Figure 9. The impact on the specific humidity retrievals is negligible. The temperature retrieval errors, however, are increased in the upper troposphere by approximately 0.08 K in the worst case scenario (Denial 3: longer than 12.82 μm excluded). The least aggressive data denial experiment (Denial 1: longer than 13.99 μm excluded) causes only a small degradation of the temperature retrievals, which is smaller than the effect of doubling the noise in the CO_2 band only. The intermediate case (Denial 2: 13.70 μm) produces a degraded temperature analysis in the upper troposphere. Given that the difference in performance between the baseline scenario and Denial 1 is smaller than the effect of any of the noise or resolution changes simulated in this study, it would appear that the noise specification above 13.99 μm can be relaxed without loss of performance.

6.4. General comments

To summarise the results of all the sensitivity experiments, Figure 10 and Figure 11 show the mean standard deviations in the lower troposphere (400-900 hPa) and upper troposphere (150-400 hPa) respectively for all three noise scenarios for both full- and half-AIRS resolution runs, as well as three CO_2 band denial runs.

The results of this study show that the two most important factors determining the quality of non-linear 1D-Var retrievals from the proposed IR sounder are the measurement noise in the H_2O band and the spectral resolution. In particular, the N2 noise case (double noise) results in a substantial increase in analysis errors in both temperature and specific humidity which should be avoided. The N3 (compromise) noise case performs well, except for upper-tropospheric temperature, where it is significantly worse than the baseline case.

The change in spectral resolution from full-AIRS resolution to half-AIRS resolution has a large effect, which for upper-tropospheric humidity is similar in magnitude to the effect of doubling the measurement noise over the entire spectrum. This effect should therefore be borne in mind when specifying the noise characteristics of the instrument. Given that the spectral resolution of the instrument is expected to be coarser than that of AIRS, it would seem advisable to try to keep the measurement noise in the H_2O band as low as possible in order to mitigate the effects of the reduced spectral resolution.

The data denial experiments described in Section 6.3 show the effect of excluding portions of the CO_2 band completely. The effect of cutting off the spectrum in the CO_2 band is particularly significant in the temperature analysis with a 12.82 µm cut-off (Denial 3). With a 13.7 µm cut-off (Denial 2) the temperature analysis is still significantly degraded, although the effect is smaller in the lower troposphere. The denial experiment with a cut-off at 13.99 µm (Denial 1) still shows some degradation in the temperature analysis in the mid- to lower troposphere.

The sensitivity studies carried out in this work have all assumed cloud-free conditions. The performance of the retrievals, particularly with differing channel selections, has not been tested in the context of cloudy radiances. Such an investigation could form the basis of a future study.

References

F. Chevallier, 2001: *Sampled databases of 60-level atmospheric profiles from the ECMWF analyses.* Eumetsat/ECMWF SAF Programme, Research Report No. 4

A. Collard, 2004: *NWP SAF Met Office 1D-Var User Manual*. NWPSAF-MO-UD-006.

C. Rodgers, 2000: *Inverse methods for atmospheric sounding.* World Scientific.

R. Saunders and P. Brunel, 2004: *RTTOV 8.5 Users' Guide*. NWPSAF-MO-UD-008.

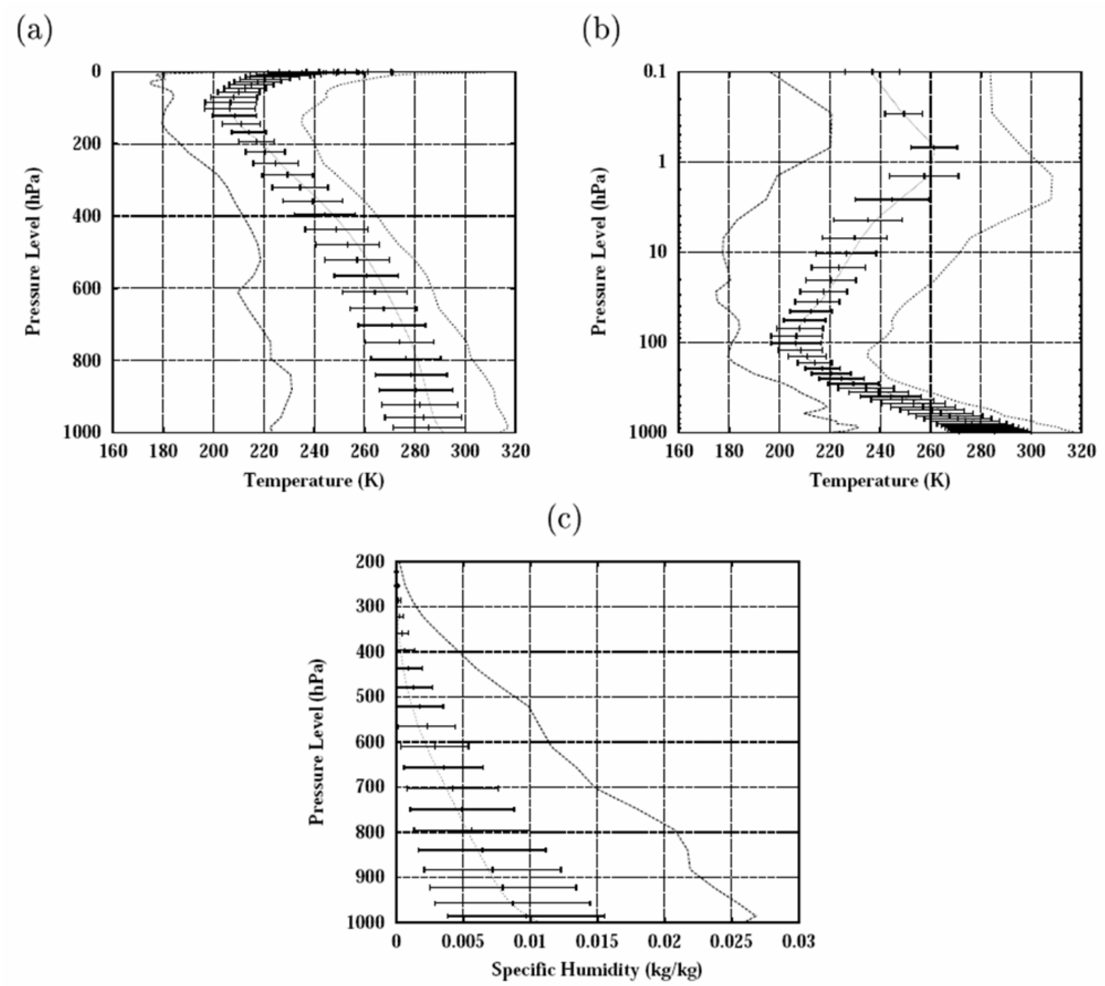


Figure 1: Variability in the ECMWF 60-level sampled dataset (from Chevallier, 2001). Solid lines indicate maxima and minima, and error bars indicate two standard deviations.

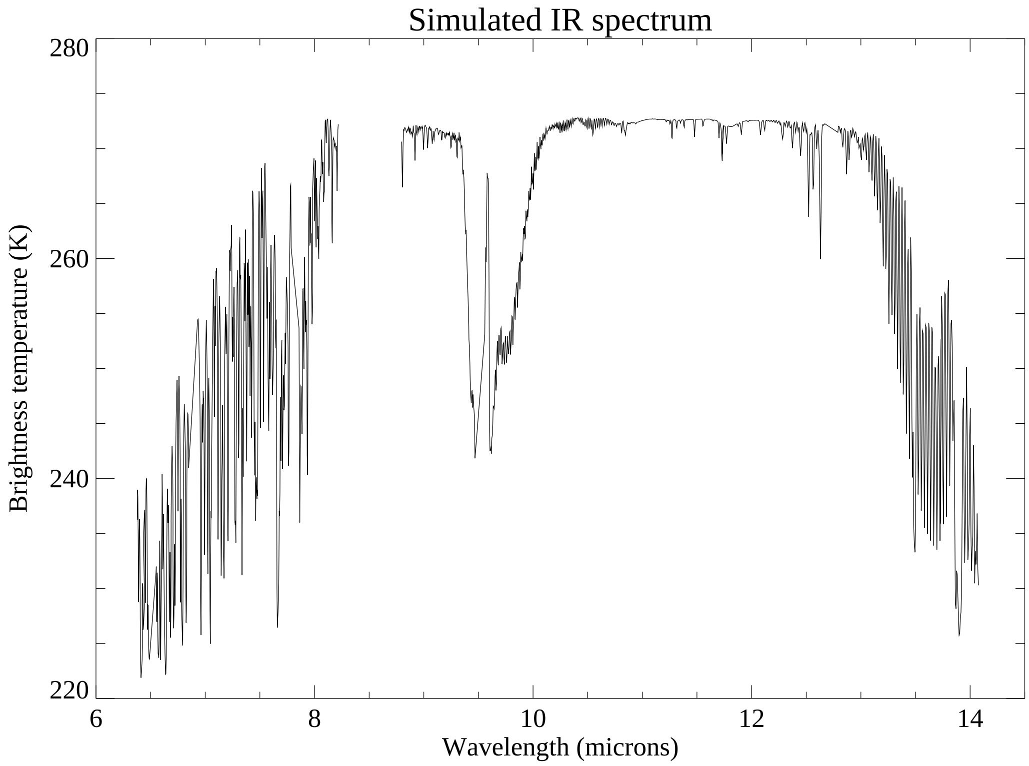


Figure 2: Example infrared spectrum as simulated for the 1D-Var study.

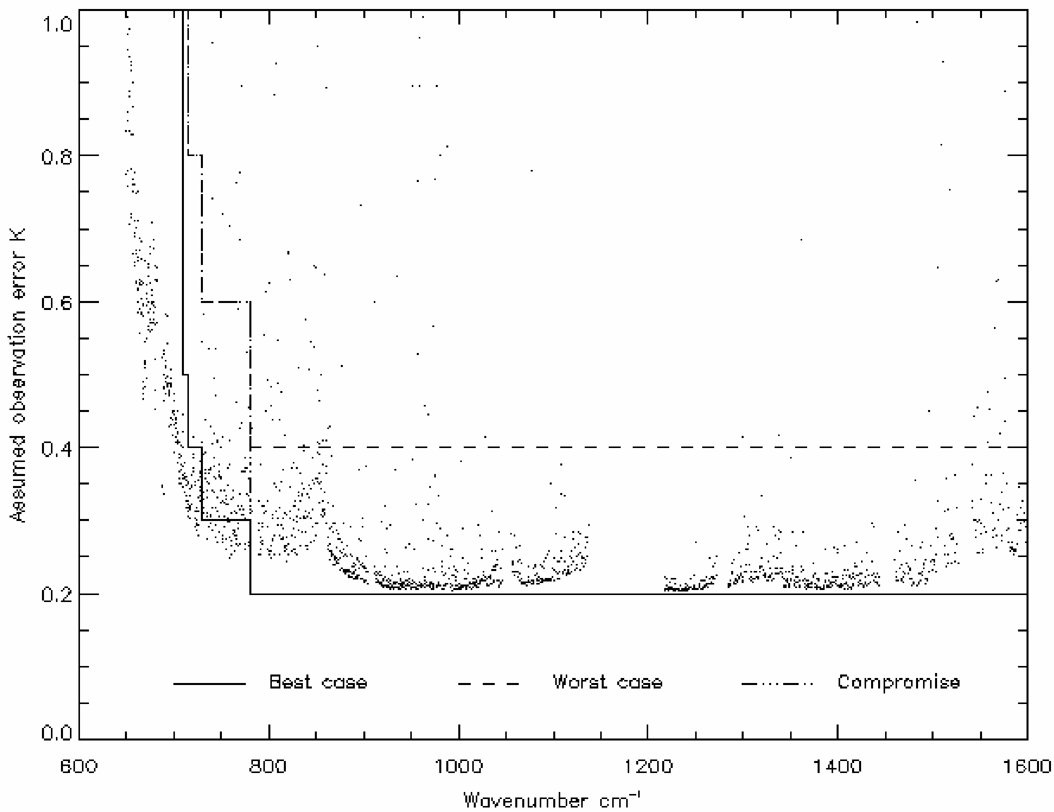


Figure 3: Measurement errors simulated in this study (at 280K scene temperature), compared with AIRS errors (dots). Solid line: best case (N1); Dashed line: worst case (N2); Dot-dashed line: compromise case (N3).

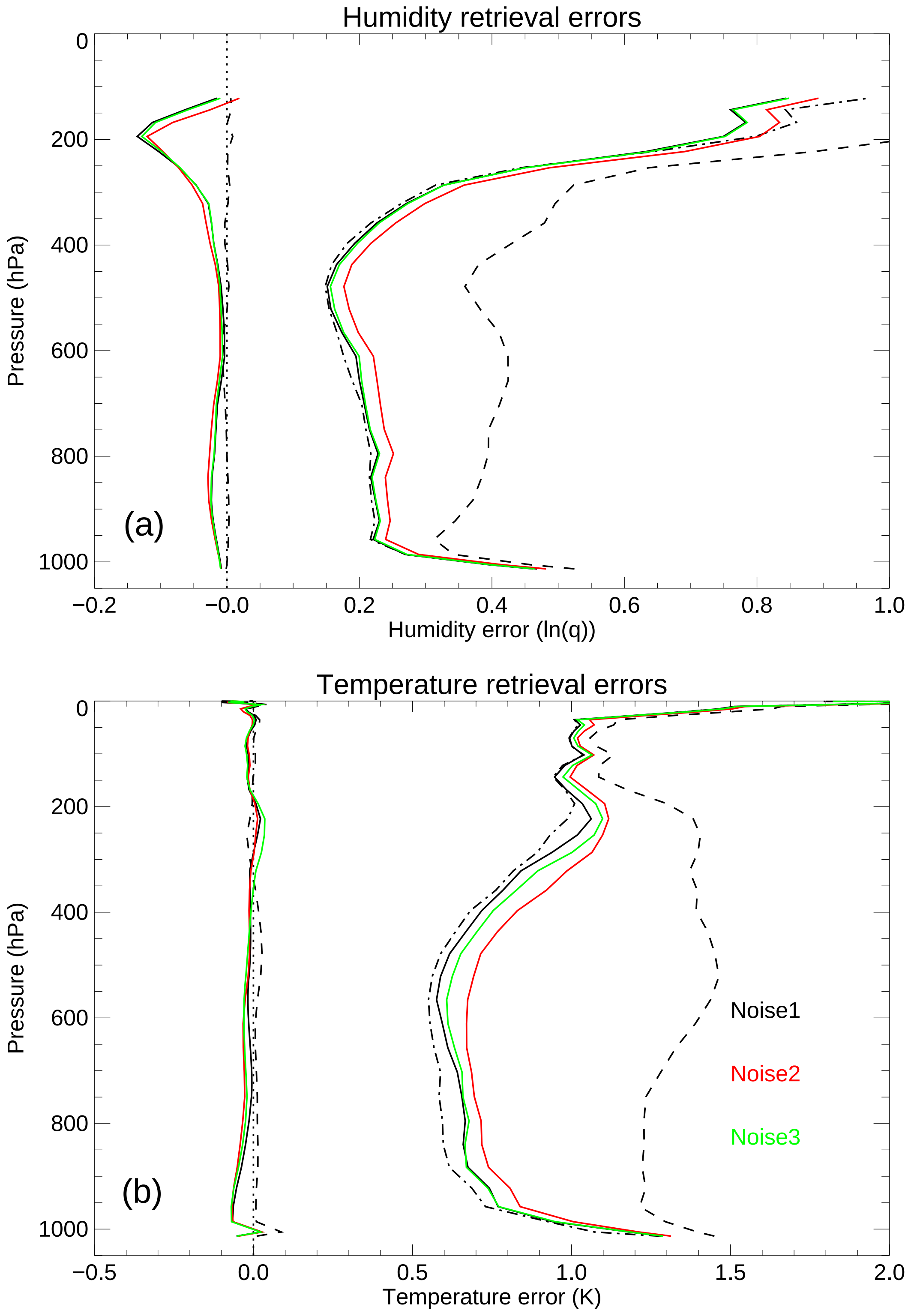


Figure 4: Retrieval errors for three noise scenarios. Left: bias; Right: standard deviation. Solid lines: retrievals; dashed lines: background; dot-dashed line: linear analysis error for N1 (best) noise case.

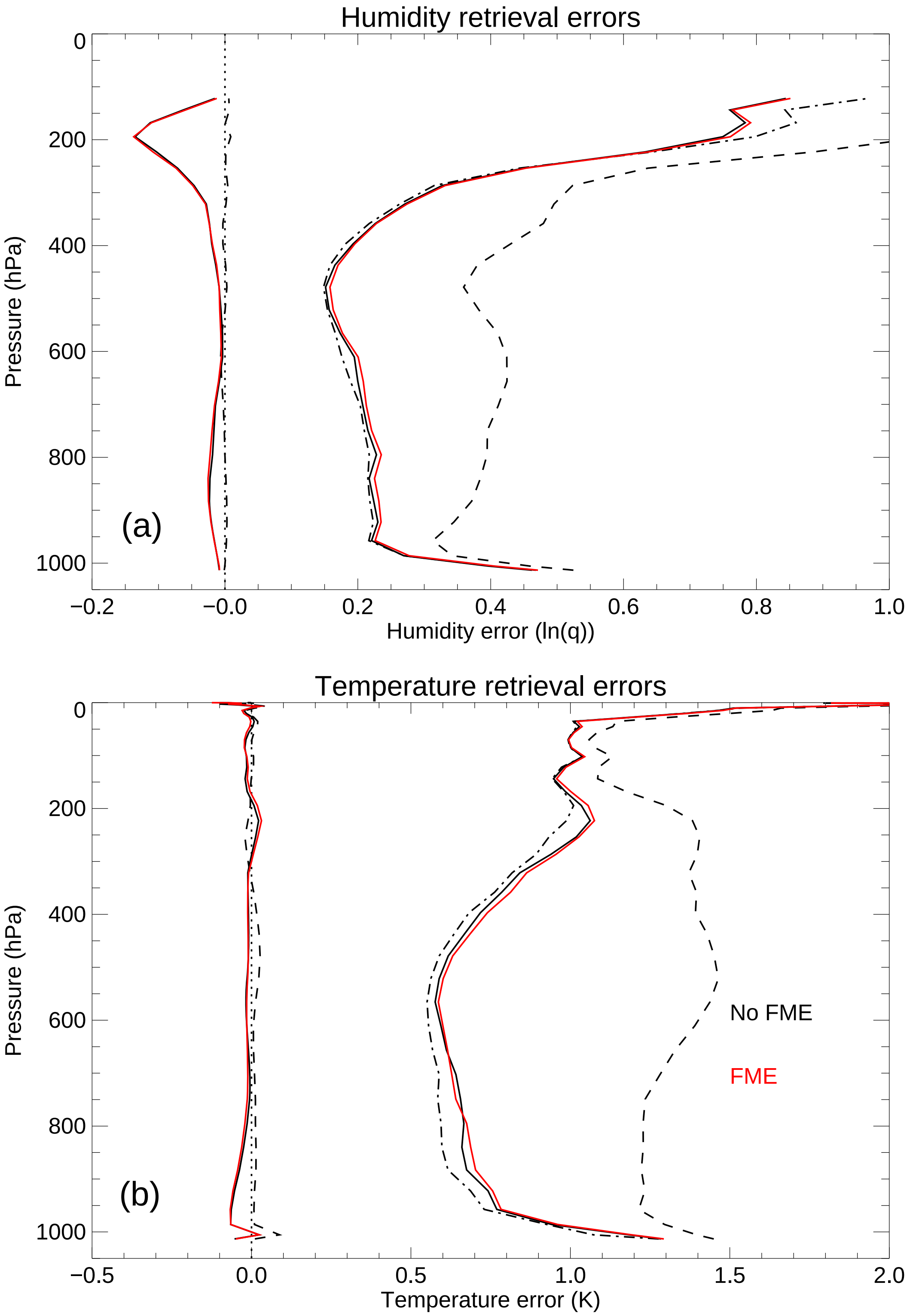


Figure 5: As Figure 4, but showing the effect of introducing a forward model error of 0.2 K (for the best (N1) noise case).

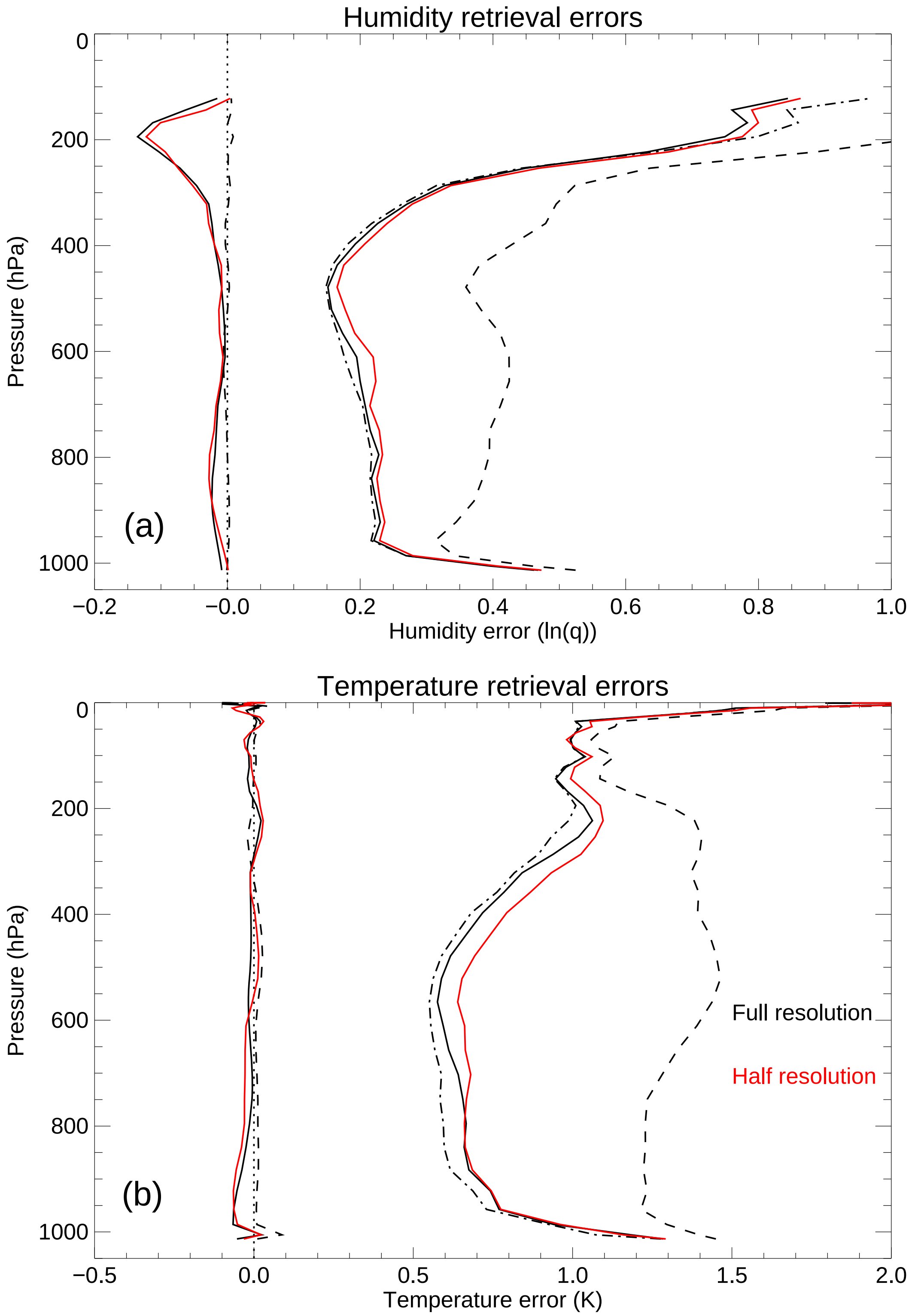


Figure 6: As Figure 4, but showing retrieval errors for full AIRS resolution (black) and half AIRS resolution (red), for the N1 (best) noise case. The dot-dashed line shows the linear analysis error expected for the N1 noise case with full resolution.

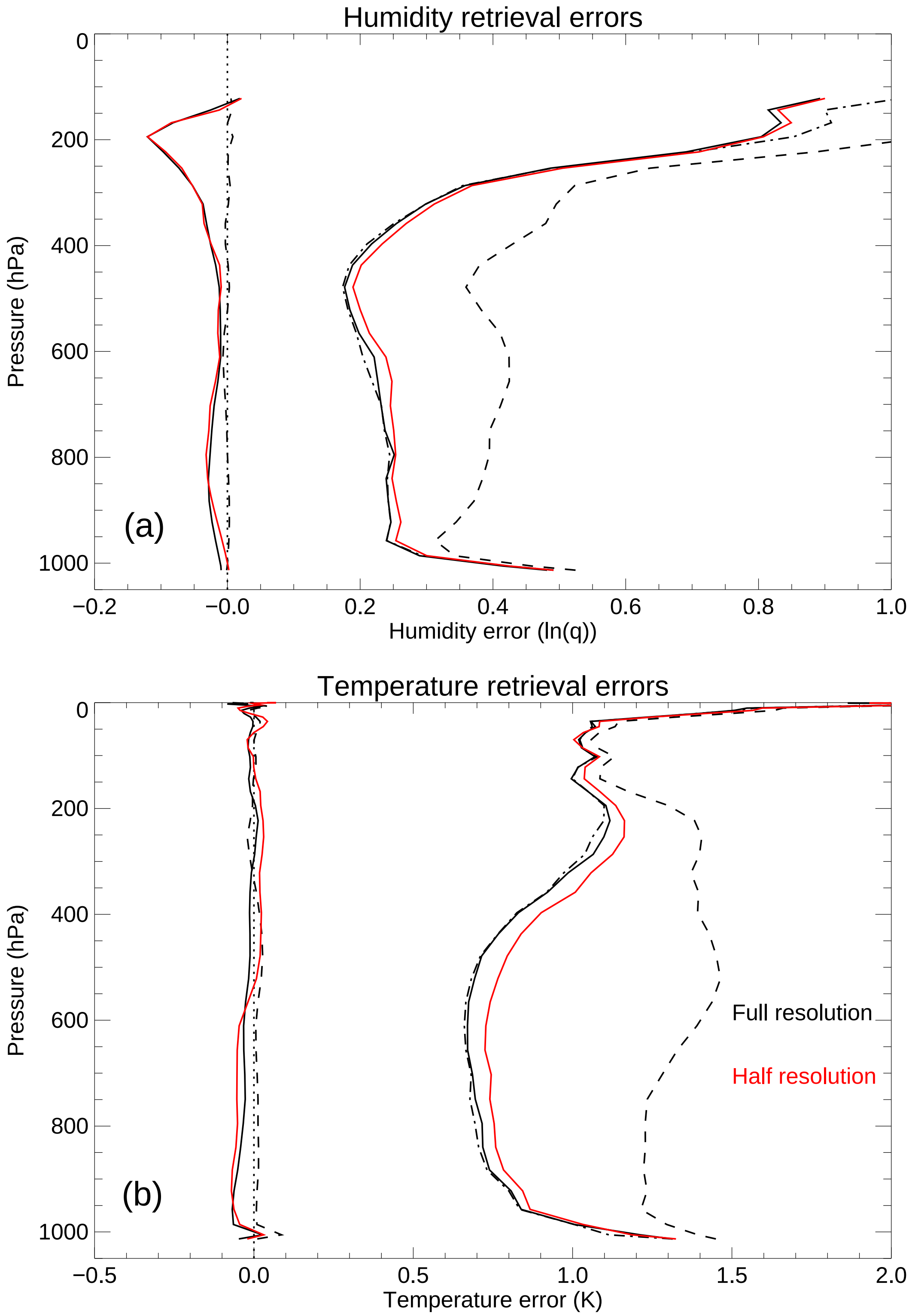


Figure 7: As Figure 4, but showing retrieval errors for full AIRS resolution (black) and half AIRS resolution (red), for the N2 (worst) noise case. The dot-dashed line shows the linear analysis error expected for the N2 noise case with full resolution.

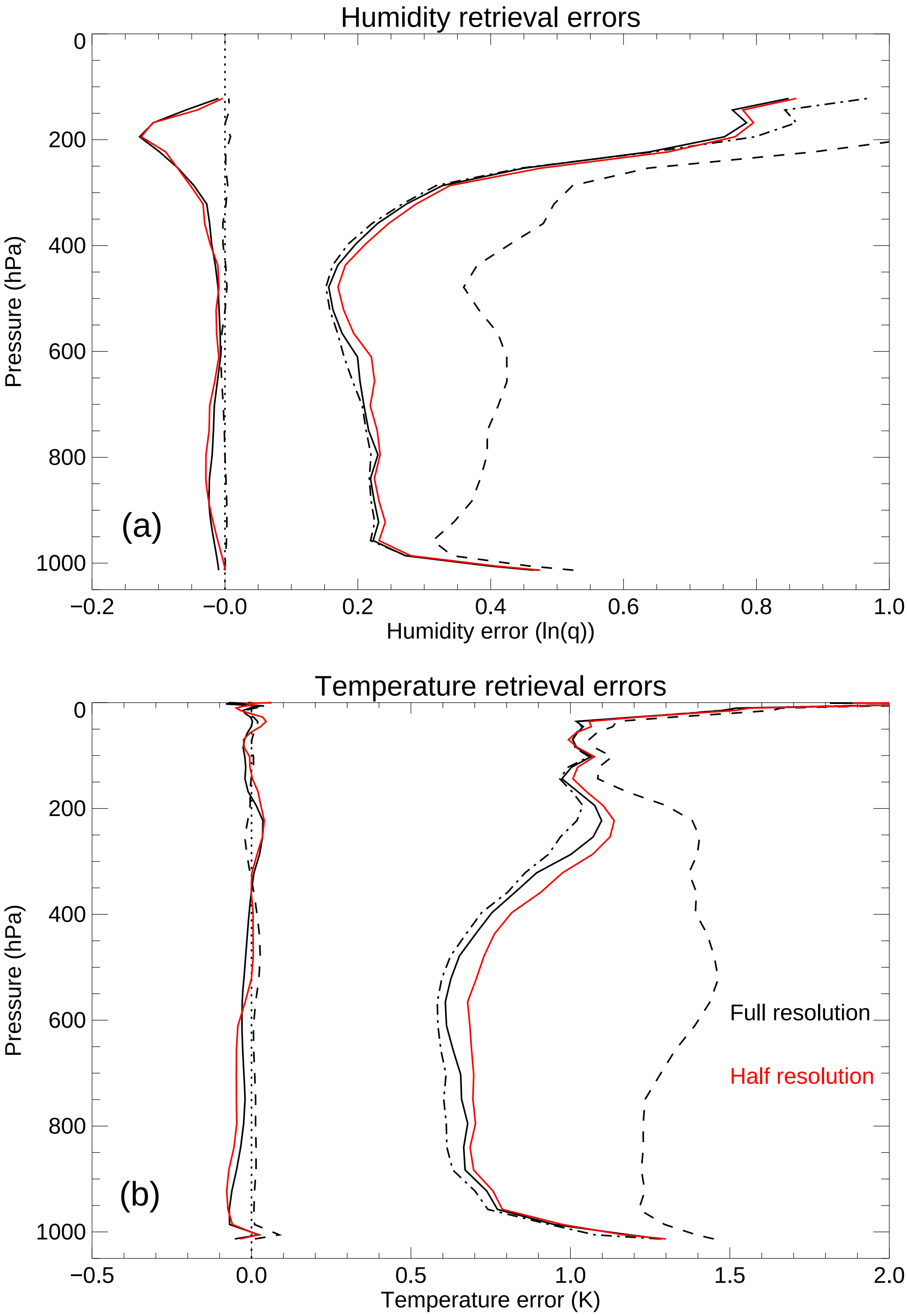


Figure 8: As Figure 4, but showing retrieval errors for full AIRS resolution (black) and half AIRS resolution (red), for the N3 (compromise) noise case. The dot-dashed line shows the linear analysis error expected for the N3 noise case with full resolution.

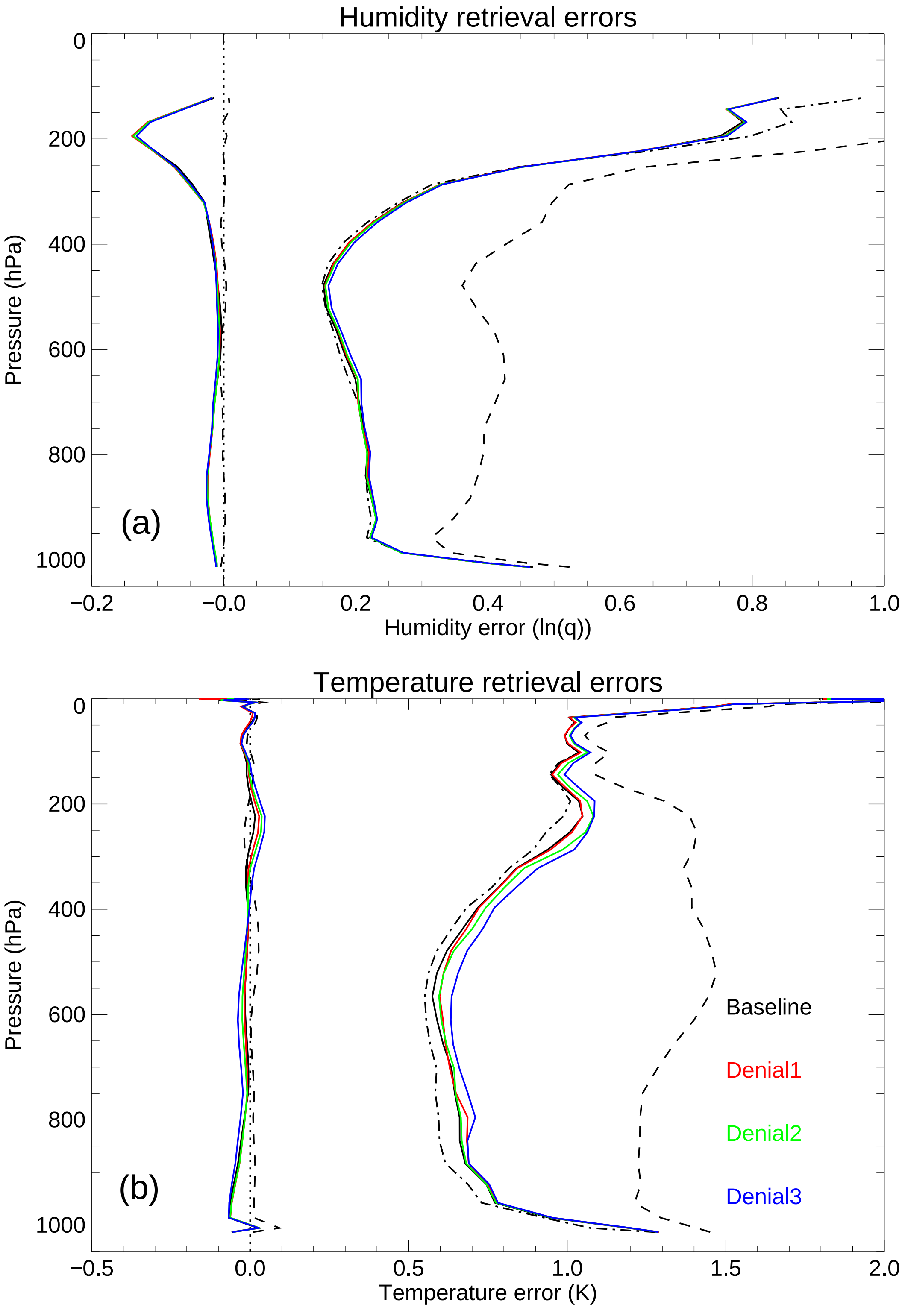


Figure 9: As Figure 4, but showing retrieval errors for the three CO_2 band data denial experiments for the N1 (best) noise case.

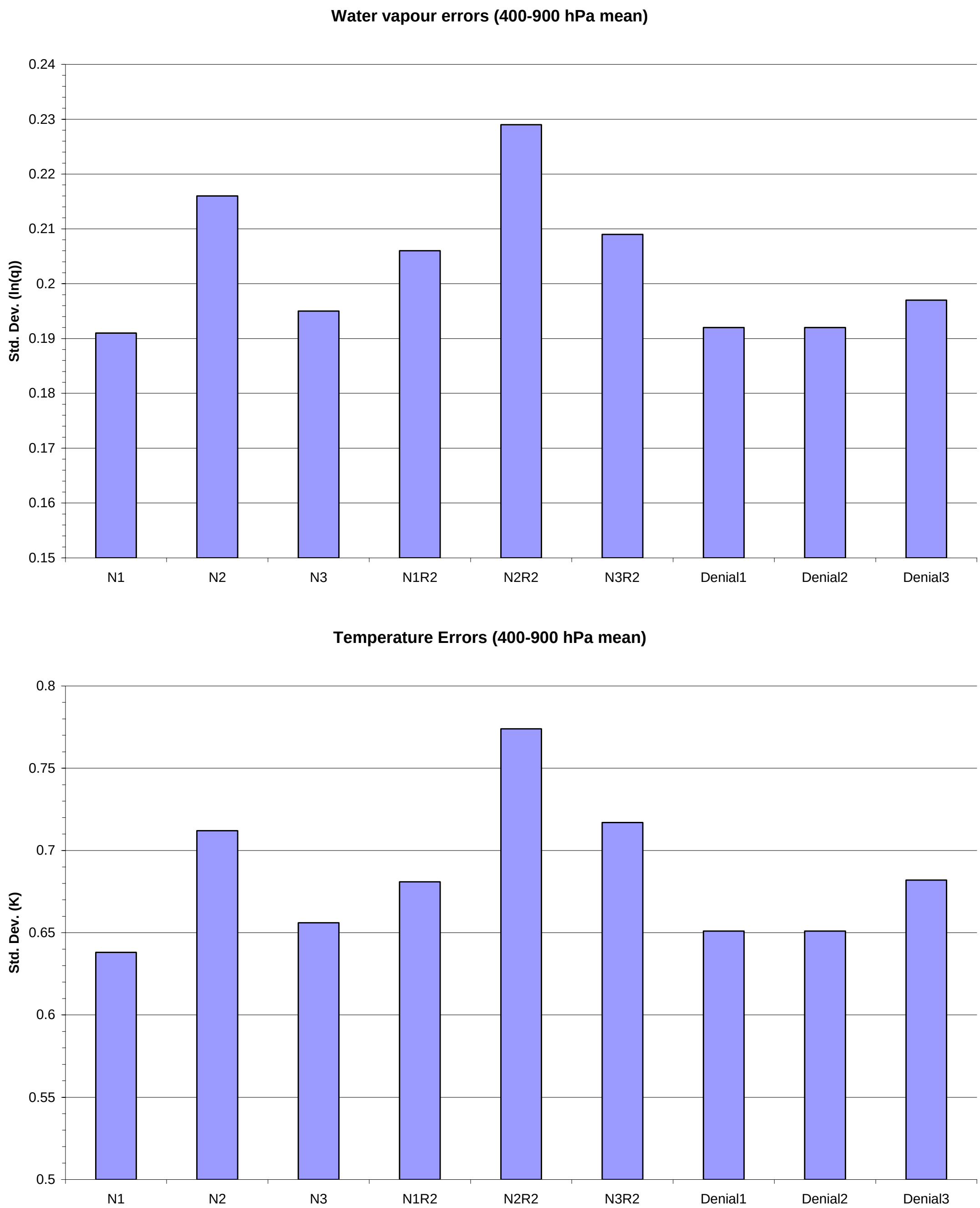
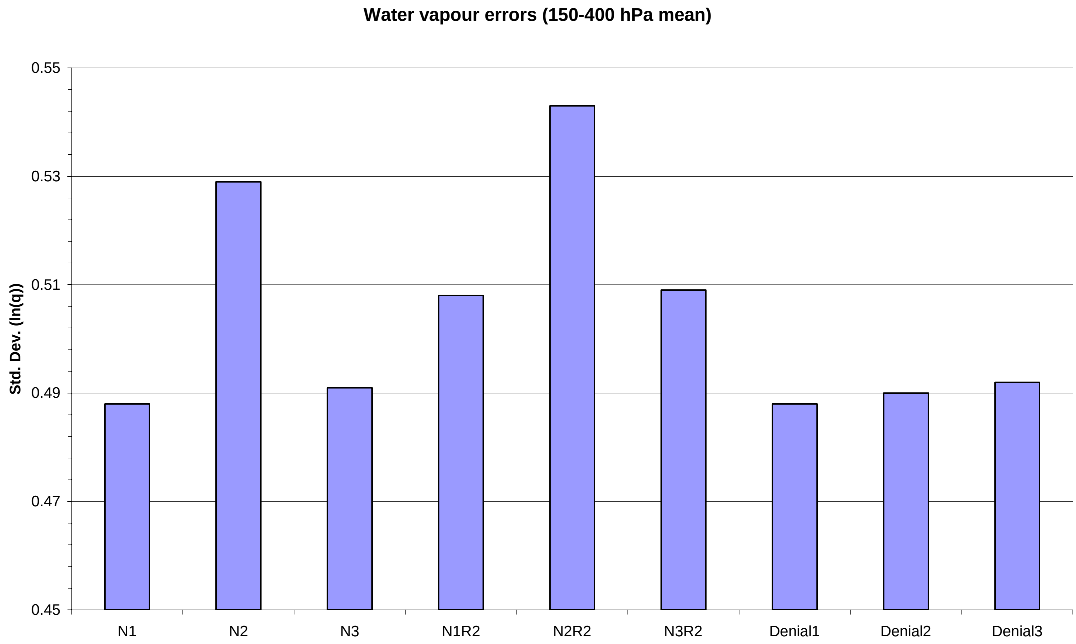


Figure 10: Mean errors between 400 and 900 hPa in water vapour and temperature retrievals for all the 1D-Var runs described in this study. N1, N2 and N3 refer to the “best”, “worst”, and “compromise” noise scenarios respectively. R1 and R2 refer to full AIRS resolution and half AIRS resolution.



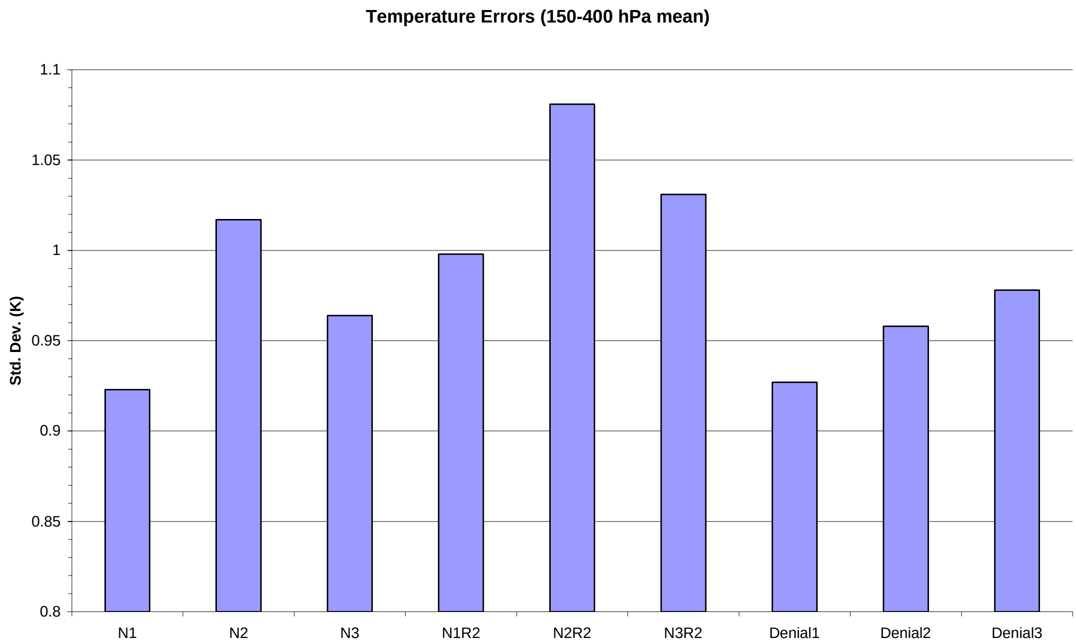


Figure 11: Mean errors between 150 and 400 hPa in water vapour and temperature retrievals for all the 1D-Var runs described in this study. N1, N2 and N3 refer to the “best”, “worst”, and “compromise” noise scenarios respectively. R1 and R2 refer to full AIRS resolution and half AIRS resolution.