Characterising extratropical near-tropopause analysis humidity biases and their radiative effects on temperature forecasts

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Characterising extratropical near-tropopause analysis humidity biases and their radiative effects on temperature forecasts

Jake Bland[1] | Suzanne Gray[1] | John Methven[1] | Richard Forbes[2]

[1]Department of Meteorology, University of Reading, Reading, UK

[2]European Centre for Medium-range Weather Forecasts, Reading, UK

Correspondence
J. Bland, Department of Meteorology, University of Reading, Reading RG6 6ET, UK.
Email: j.s.bland@pgr.reading.ac.uk

Funding information
Natural Environment Research Council, Grant/Award Number: NE/L002566/1

Abstract

A cold bias in the extratropical lowermost stratosphere in forecasts is one of the most prominent systematic temperature errors in numerical weather prediction models. Hypothesized causes of this bias include radiative effects from a collocated moist bias in model analyses. Such biases would be expected to affect extratropical dynamics and result in the misrepresentation of wave propagation at tropopause level. Here the extent to which these humidity and temperature biases are connected is quantified. Observations from radiosondes are compared to operational analyses and forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) and Met Office Unified Model (MetUM) to determine the magnitude and vertical structure of these biases. Both operational models over-estimate lowermost stratospheric specific humidity, with a maximum moist bias around 1 km above the tropopause where humidities are around 170% of the observed values on average. This moist bias is already present in the initial conditions and changes little in forecasts over the first five days. Though temperatures are represented well in the analyses, the IFS forecasts anomalously cool in the lower stratosphere, relative to verifying radiosonde observations, by $0.2 \, \text{K day}^{-1}$. The IFS single column model is used to show this temperature change can be attributed to increased long-wave radiative cooling due to the lowermost stratospheric moist bias in the initial conditions. However, the MetUM temperature biases cannot be entirely attributed to the moist bias, and another significant factor must be present. These results highlight the importance of improving the humidity analysis to reduce the extratropical lowermost stratospheric cold bias in forecast models and the need to understand and mitigate the causes of the moist bias in these models.

KEYWORDS
humidity observations, long-wave cooling, lowermost stratosphere, radiosondes, single-column model

1 | INTRODUCTION

The representation of specific humidity near the tropopause in numerical models has been shown to be important for the accuracy of medium-range forecasts and climate integrations. Modelling studies have demonstrated that both stratospheric and tropospheric temperatures in climate models are sensitive to stratospheric water vapour (Smith *et al.*, 2001; Solomon *et al.*, 2010). Many of these studies have been motivated by an observed trend of increasing stratospheric water vapour in the late 20th century, and show that this results in enhanced cooling of the lower stratosphere (Forster and Shine, 1999; Maycock *et al.*, 2011). If the water vapour concentration in the lowermost stratosphere is increased, this will increase both the emittance of the stratosphere and absorptance of upwelling radiation from the troposphere. In the strong infrared absorption bands of water vapour, the increased emission dominates over the increased absorption (Shine and Myhre, 2020). Therefore an increase in water vapour in the lowermost stratosphere would lead to an increase in emission from water vapour, which would lower the lower-stratospheric temperature required for the outgoing radiation to balance the incoming radiation in equilibrium (Maycock *et al.*, 2011). By increasing stratospheric water vapour in general circulation models (Joshi *et al.*, 2006; Maycock *et al.*, 2013) or imposing stratospheric cooling to mimic the temperature response to increased stratospheric water vapour (Tandon *et al.*, 2011), it has also been shown that such changes can lead to a poleward shift in the jets and storm tracks, a strengthening of the jets and other changes to the atmospheric circulation in models. The aim of this paper is to characterise lowermost stratosphere humidity biases in atmospheric analyses and their impact on temperature biases in numerical weather prediction forecasts.

Radiative transfer near the tropopause also plays an active role in maintaining tropopause sharpness with effects on large-scale dynamics. The abrupt drop in specific humidity immediately above the extratropical tropopause results in a peak in long-wave cooling from the troposphere immediately below the tropopause (due to lack of water vapour in the layers above). It has been shown using radiative transfer modelling (Randel *et al.*, 2007) and further supported by observational analysis (Hegglin *et al.*, 2009) that lower-stratospheric water vapour plays an important role in maintaining the region of enhanced static stability immediately above the tropopause, often called the tropopause inversion layer (TIL; Birner 2006). When considering large-scale dynamics, the tropopause is typically defined as an iso-surface of potential vorticity (PV) and the value of 2 PVU (PV units) is often used in midlatitudes (Hoskins and James, 2014). PV is a measure of rotation and stratification in the atmosphere and the sharp change in static stability at the tropopause is associated with a strong PV gradient and a marked change in wind shear. Since PV is materially conserved for adiabatic frictionless flows, this definition highlights that the tropopause is approximately a material surface; this cannot be deduced from the temperature profile alone. The long-wave cooling from the tropopause level results in a dipole of diabatically generated PV that is positive above the tropopause and negative below, and so acts to sharpen the PV gradient (Forster and Wirth, 2000; Chagnon *et al.*, 2013; Saffin *et al.*, 2017) – an alternative description for the formation of the TIL. Additional water vapour in the lowermost stratosphere is expected to weaken the diabatic PV dipole and hence tends to reduce the PV contrast across the tropopause zone. Gray *et al.* (2014) found evidence for a marked decrease in PV gradient with lead time in global forecasts (from several operational centres) although they did not quantify the processes contributing to this decline. The unrealistic decline in PV gradient in forecasts has ramifications for Rossby waves propagating along the tropopause. Theoretical considerations have shown that smoothing PV gradients reduces Rossby wave phase speed (Harvey *et al.*, 2016) and is expected to reduce the amplitude of large-scale jet meanders due to excessive PV filamentation and flux of wave activity away from the jet core (Harvey *et al.*, 2018). In summary, the representation of the humidity contrast across the tropopause is expected to affect radiative heating profiles, tropopause gradients in temperature and wind, and large-scale dynamics.

As stratospheric water vapour impacts atmospheric radiative balance, its representation in simulations has been evaluated. It has been known for at least 20 years that atmospheric model analyses, re-analyses and forecasts are typically moister than observed in the extratropical lower stratosphere (Pope *et al.*, 2001). This bias has been shown through comparisons to many different observational datasets as summarised in Table 1. For the same reasons that a trend of increasing stratospheric water vapour would lead to a cooling of the lower stratosphere, one would expect radiative effects resulting from a moist model bias in the lowermost stratosphere in the analysis to lead to a cold bias in the extratropical lowermost stratosphere in forecasts (Stenke *et al.*, 2008; Diamantakis and Flemming, 2014; Shepherd *et al.*, 2018). Direct measurements of temperature from radiosondes and aircraft as well as indirect measurements from satellite radiances and radio occultation are assimilated into both models which constrain the temperature in the analyses. Dyroff *et al.* (2015) and Carminati *et al.* (2019) show the mean temperature errors of the ECMWF and MetUM analysis in the extratropical lowermost stratosphere to be within a few

TABLE 1 Summary of model moist biases in the upper troposphere and lower stratosphere found in other studies

Reference	Observation Type	Model/Analysis	Detail of humidity bias
Stenke *et al.* (2008)	UARS HALOE (Satellite)	ECHAM4.L39	Model moist bias by factor of 3–5 compared to observations in the extratropical lowermost stratosphere.
Van Thien *et al.* (2010)	Aura MLS (Satellite)	GFS and NAM model	Both models are more moist than observations over North America and nearby ocean regions between 150 and 250 hPa.
Oikonomou and O'Neill (2006)	MOZAIC (Aircraft)	ERA-40	ERA-40 has up to over 60% higher mixing ratios than observations in the Northern Hemisphere mid- and high-latitude lower stratosphere, as a fraction of the observed values. The wet bias is larger in summer than in winter.
Feist *et al.* (2007)	AMSOS (Aircraft)	ERA-40	ERA-40 is moister than observations in the Northern Hemisphere lower stratosphere.
Kunz *et al.* (2014)	FISH (Aircraft)	ERA-Interim and ECMWF operational analyses	Analyses and Reanalyses underestimate high water vapour mixing ratios in the upper troposphere and overestimate low mixing ratios in the lower stratosphere.
Dyroff *et al.* (2015)	CARIBIC (Aircraft)	ECMWF operational analysis and short (< 24 hr) forecasts	ECMWF has a moist bias in the lowermost stratosphere. This moisture excess is largest in summer with a maximum median difference around 2 km above the tropopause with the model humidity ≈ 220% of observed values, and smallest in winter with a maximum around 3 km above the tropopause with model humidity ≈ 150% of observed values.
Woiwode *et al.* (2020)	GLORIA (Aircraft)	ECMWF operational analysis and short (< 12 hr) forecasts	Systematic moist bias in the polar lowermost stratosphere with model humidities on average ≈ 150% of the observed values.

tenths of a Kelvin. Although the subsequent growth of a cold bias in the first days of the forecast in this region is seen in operational verification, the authors are not aware of documentation of this time-range in the published literature. However, the development of the cold bias in the longer forecast range and climate of each model has been described. For example, Gates *et al.* (1999) showed a cold bias in the lower polar stratosphere in the Atmospheric Model Intercomparison Project (AMIP) ensemble compared to the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-15 reanalysis. Extratropical lower-stratospheric cold biases with respect to ERA-Interim reanalyses of up to 5 K were found more recently in 20-year AMIP-type simulations with the Met Office Unified Model (MetUM; Hardiman *et al.*, 2015; Oh *et al.*, 2018) and in multiple 1-year free-running simulations with the ECMWF Integrated Forecasting System (IFS; Shepherd *et al.*, 2018). The ECMWF IFS was also

shown to have a more severe cold bias in forecasts in the summer hemisphere than the winter hemisphere, which may be related to the larger moist bias in ECMWF analyses in summer than in winter found by Dyroff *et al.* (2015) through the additional radiative cooling this would cause.

The aims of this study are firstly to identify and characterise humidity and temperature biases in the upper troposphere and lowermost stratosphere (UTLS) in the IFS and MetUM, and secondly, to determine the extent to which these temperature biases can be attributed to the presence of the diagnosed moist bias and explore the mechanism by which the moist bias and temperature biases may be causally related. The first aim of characterising any biases is a necessary step to determining their sources in weather forecasts. We use radiosonde observations obtained predominantly over the eastern North Atlantic and Western Europe for the two-month period of September and October 2016 to evaluate differences in
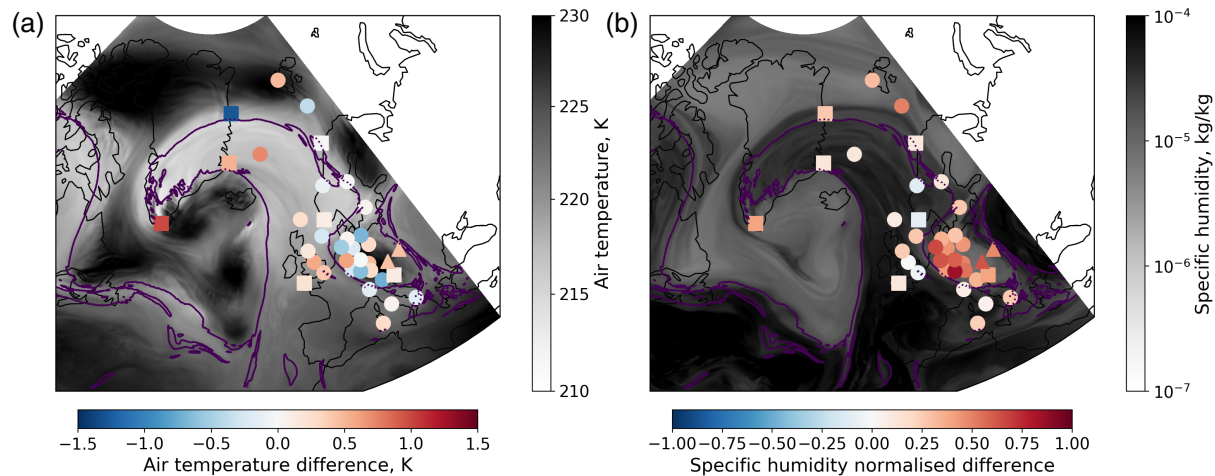
FIGURE 1 Maps of (a) air temperature and (b) specific humidity from IFS operational analyses on 03 October 2016 1200 UTC at 250 hPa (greyscale shading) with the location of the dynamical tropopause shown by the 2 PVU contour (purple). Within the coloured shapes are (a) the difference between the IFS temperature and the temperature at locations as measured by radiosondes launched from these locations at the same time and pressure level, (b) the normalised difference of specific humidity between the model and the observations as defined in Section 3.5. Squares indicate sites using the RS41 radiosonde, circles those using RS92, and triangles those using a combination of the two

specific humidity and temperature between observations and IFS and MetUM operational analyses and forecasts, with a focus on tropopause relative vertical structure and the relationship between biases. The main benefit of radiosonde data over satellite data or *in situ* aircraft observations made along flight tracks which we take advantage of for this study is the high vertical resolution of the observations. This better facilitates the investigation of the vertical structures of any biases in the UTLS, and allows calculation of the observed tropopause altitude for tropopause-relative compositing, and evaluation of the model representation of the tropopause altitude. The collocated temperature and humidity measurements are also important for determining any connection between such biases. Furthermore, as radiosonde humidity observations are not assimilated above the tropopause, the lower-stratospheric humidity observations provide an independent dataset against which to assess analyses. To address the second aim of understanding the relationship between biases, the ECMWF Single-Column Model (SCM) is used to investigate the radiative impact of the systematic lowermost stratosphere specific humidity biases in analyses, and to determine what proportion of the systematic temperature biases in forecasts can be attributed to this.

In Section 2 we describe the radiosonde data used in this comparison, the two numerical weather prediction models and the SCM. We then outline the methods used for the comparison in Section 3. In Section 4 the results of the comparison of model data to radiosonde observations are presented, followed by the results from the SCM experiments and a discussion of changes at

tropopause level resultant from lowermost stratosphere humidity differences in Section 5. The main conclusions are then summarised in Section 6.

2 | DATA

2.1 | Radiosondes

In this study we use data from 3204 radiosondes which were launched from 40 sites indicated in Figure 1 over the North Atlantic region (38°–80°N, 50°W–24°E) in September and October of 2016. Of these, 2602 are of the type Vaisala RS92 (Vaisala, 2013) and 602 are of the newer type Vaisala RS41 (Vaisala, 2017b). From 33 sites radiosondes were typically released twice per day, and from the rest once per day. The radiosonde ascents are mostly operational launches, but additional launches were also made for the NAWDEX (North Atlantic Waveguide and Downstream Impact EXperiment) field campaign: a project with the aims of exploring the impact of diabatic processes on the jet stream and midlatitude weather systems (Schäfler *et al.*, 2018). The radiosondes used either reported measurements every 2 s, which corresponds to approximately every 10 m, or at significant levels (Ingleby *et al.*, 2016).

The resolution and total uncertainty of the temperature measurements made by the RS92 radiosondes are 0.1° and 0.5°C, respectively. The resolution of the relative humidity (RH) data is 1% RH and the total uncertainty is 5% RH for temperatures >−60°C (Vaisala, 2013). Total uncertainty

here refers to a two-standard-deviation confidence level, including repeatability and effects due to measurement conditions, response times and measurement electronics. The RS41 radiosondes have a resolution of 0.01°C and a combined uncertainty of 0.4°C for temperature measurements, and a resolution of 0.1% RH and combined uncertainty of 4% RH for humidity measurements (Vaisala, 2017a). For both radiosonde types the reproducibility in soundings is 2% RH. These figures are taken from Vaisala datasheets, and further information on the measurement performance can be found in Vaisala (2017a). The WMO intercomparison of radiosonde systems (Nash *et al.*, 2010) shows that the Vaisala RS92 radiosonde performs well in comparison to a Cryogenic Frostpoint Hygrometer (CFH) for humidity measurements, including in the upper troposphere and lower stratosphere, showing that systematic errors are less than 2% RH. The improvement in performance of the RS92 compared to earlier studies is due to improved sensor coating and correction algorithms for solar radiation and time lag, removing previously found biases (Vaisala, 2020; Wang *et al.*, 2013). Comparison studies show good agreement in both the temperature and humidity measurements between the RS92 and RS41 radiosondes (Jauhiainen *et al.*, 2014). Although the differences in measurements are small, the RS41 demonstrates a better precision and a reduced sensitivity to solar heating (Edwards *et al.*, 2014; Motl, 2014; Jensen *et al.*, 2016). In terms of bias between the two instruments, it can be seen in intercomparison studies (Edwards *et al.*, 2014; Jauhiainen *et al.*, 2014; Vaisala, 2014; Dirksen *et al.*, 2020) that the RS92 is < 1.5% RH drier than the RS41 in the upper troposphere and < 1% RH moister in the lower stratosphere, which is a very close agreement given the uncertainties of 4% RH and 5% RH for the two sonde types in these measurements.

Although the measured quantity is relative humidity, the humidity variable reported by radiosondes is dew-point temperature. The resolution of these measurements is 0.01°C, as for temperature. The humidity quantity which we mainly consider in this article is specific humidity, the calculation of which from the dew-point temperature is detailed in Section 3. To provide an indication of how large the measurement uncertainties and biases detailed above are as a fraction of the mean specific humidity at a given altitude, the measurement uncertainty of 5% RH as given in the RS92 datasheet corresponds to around 5–10% uncertainty in mean specific humidity in the troposphere. This increases from the top of the tropopause to around 50–100% of the mean values 2 km above the tropopause. In the lowermost stratosphere, this measurement uncertainty is much larger than other sources of uncertainty such as uncertainty in temperature and precision of the dew-point temperature. The mean

relative humidity more than 2 km above the tropopause is below 5% RH, and therefore we acknowledge that humidities measured at higher altitudes are associated with very large uncertainty. It is also noted in Edwards *et al.* (2014) and Nash *et al.* (2010) that there are diurnal differences in the performance of the RS92 radiosondes. By comparing data from radiosondes released at 1200 UTC to those released at 0000 UTC in our dataset, we find that the RS41 radiosondes exhibit negligible day/night differences, but that the RS92 radiosondes report slightly higher humidity during the day, with the difference being everywhere less than 3% of the mean specific humidity value in the troposphere, and less than 5% in the stratosphere. Measurement uncertainties in relation to the biases we observe are discussed in more detail in Section 4.1.2.

As noted above, the radiosonde data from this period are transmitted to the WMO Global Telecommunications System (GTS) in one of two different formats. Twenty six of the sites report the measurement data recorded every 2 seconds, giving a vertical spacing of approximately 10 m throughout the ascent. For the other eighteen sites, radiosonde data are sub-sampled and transmitted only for significant levels: a set of mandatory pressure levels, in addition to altitudes chosen on each profile where there is a marked change in the gradient of the temperature or humidity . The profile obtained by linear interpolation between significant levels is, by design, very similar to the raw high resolution profile and the significant levels are only used to reduce data transmission from remote sites.

To make the observed data comparable to data from the models in terms of smoothness, we apply a Gaussian kernel smoothing filter to the profiles, of half-width 200 m (Harvey *et al.*, 2020). This smooths the observed data to a resolution similar to that of the models. The agreement between the altitudes of the tropopause from observations and the models were compared for different smoothing Gaussian half-widths. Half-widths of greater than 200 m gave no improvement on the agreement and failed to resolve features of interest, while those smaller than 200 m resolved features too finely, giving an increased median difference in calculated tropopause height between model and the observations (not shown). For the radiosonde data on significant levels, the radiosonde data were first linearly interpolated between observation points to a 10 m grid before smoothing to make them comparable to the radiosonde data from the high-resolution sites.

2.2 | Models

Vertical profiles from the radiosondes are compared to profiles taken from the operational IFS and MetUM forecast models. The ECMWF analysis and forecast data

are interpolated to a 0.125° × 0.125° latitude–longitude grid. The operational version in autumn of 2016 of the ECMWF's high-resolution atmospheric model was IFS cycle 41r2 (ECMWF, 2016) with horizontal resolution TCo1279 (~9 km grid spacing; Malardel *et al.*, 2016) and 137 vertical levels. The mean vertical model level spacing at the tropopause in the midlatitudes is approximately 300 m. The Met Office analyses and forecasts from the NAWDEX period were produced using the MetUM version 10.2 in the operational global configuration GA6.1 (Walters *et al.*, 2017), with a horizontal resolution of N768 (~17 km grid spacing) and 70 vertical levels. Data are linearly interpolated in the horizontal to radiosonde release sites. The vertical model level spacing at the tropopause in the midlatitudes is approximately 550 m (Schäfler *et al.*, 2020). Radiosonde data are compared to model data from the nearest six-hourly analysis or forecast. Operational radiosondes are typically launched 45 min before their nominal report time, so at typical ascent rates the radiosonde is expected to be close to the tropopause at 0000, 0600, 1200 or 1800 UTC.

As noted in the previous section, increments to model humidity from radiosondes are not used above the tropopause in data assimilation in the IFS (Ingleby, 2017), and in the MetUM they are not used above the 5 PVU surface, or between 2.5 and 5 PVU if observed humidity values fall outside of the climatological range of 1–3 ppm and relative humidity < 10% RH, between 100 and 400 hPa (Ingleby *et al.*, 2012). There are several difficulties with assimilating humidity data in the stratosphere, including the lack of available near-real-time high-quality observations with sufficient vertical resolution and global coverage. Even with radiosonde or aircraft data, there are difficulties allowing humidity increments into the stratosphere due to the sharp humidity gradient at the extratropical tropopause, as small displacements can lead to large differences between observed and background humidities (Bannister *et al.*, 2020). A similar problem is found for the large gradients associated with boundary-layer inversions where vertical positional errors can lead to degradation in the analysis (Fowler *et al.*, 2012). Furthermore, depending on the humidity variable used, allowing humidity increments in the absence of observations can lead to a moistening of the lower stratosphere as the assimilation corrects for the cool bias, due to the correlation between temperature and humidity (Dee and Da Silva, 2003).

The IFS SCM represents the physical processes in a vertical column for a single grid-point in the horizontal. We use it here to isolate the changes in the response of these physical processes to changes in the initial vertical profiles of variables from effects due to the larger-scale dynamics. For the SCM experiments we use version 43r3 of the IFS, which is a later version than used

for the full model simulations, but which has very similar lower-stratospheric temperature errors. The SCM is run with the same 137 vertical levels as the full model. The physical processes included are as detailed in the IFS 43r3 documentation (ECMWF, 2017). Further detail on how we forced the SCM for these experiments is provided in Section 5.1, and more information about the IFS SCM can be found at, for example, Carver (2019).

3 | METHODS

In this section we outline the methods used for the calculation of specific humidity from radiosonde-reported dew-point temperature (so that these values can be compared to the model output), the method used to identify the tropopause and the tropopause-relative coordinates that are used throughout the paper to calculate the mean properties of the extratropical lowermost stratosphere. We then explain how we ensure that the tropopause-relative comparisons we make are appropriate, and the metric used for the evaluation of the specific humidity biases.

3.1 | Calculation of specific humidity from radiosonde ascents

Before the smoothing filter is applied to the radiosonde data, specific humidity is calculated from the dew-point temperature reported by the operational radiosonde data processing system. The saturation vapour pressure, e, is first calculated from the dew point temperature using the Sonntag numerical approximation (Sonntag, 1990; 1994), chosen because it is used in the humidity observation operators in the data assimilation for both the IFS and Met Office (Ingleby and Edwards, 2015; Haiden *et al.*, 2016):

$$\ln\left(\frac{e}{100}\right) = -\frac{6096.9385}{T_d}$$
$$+ 16.635794 - 2.711193 \times 10^{-2} T_d$$
$$+ 1.673952 \times 10^{-5} T_d^2 + 2.433502 \ln(T_d),$$

where e is in Pa and T_d is dew-point temperature in Kelvin. The specific humidity, q, in kg kg^{-1} is subsequently calculated as

$$q = \frac{\epsilon e}{\max(p, e) - (1 - \epsilon)e},$$

where p is pressure in Pa and ϵ is the ratio between the specific gas constant for dry air and the specific gas constant for water vapour, R_d/R_v, equivalent to the ratio of the molar masses of water vapour and dry air. As q is unitless, for the remainder of the paper units of q will not

be included in the text. The use of the maximum function here is to restrict the maximum value of q to 1 if, for any reason, the partial pressure, e, is calculated as being larger than the air pressure, p.

3.2 | The thermal tropopause

As we are concerned mainly with the upper troposphere/lower stratosphere (UTLS) region, we also need to calculate the altitude of the tropopause. The thermal tropopause is found using the World Meteorological Organisation definition (WMO, 1957): "The lowest level at which the lapse rate decreases to $2\,\mathrm{K\cdot km^{-1}}$ or less, provided that the average lapse rate between this level and all higher levels within 2 km does not exceed $2\,\mathrm{K\cdot km^{-1}}$", with an additional requirement that the mean specific humidity in the 1 km layer above the tropopause should be less than 4×10^{-5}. This latter threshold is applied to reduce the number of cases when the tropopause is found as a lower-level inversion, and is chosen to be sufficiently high that values of specific humidity in the stratosphere are always less than this.

3.3 | Tropopause-relative coordinates

Once the altitude of the thermal tropopause has been calculated, using the tropopause altitude identified from the radiosonde observations, z_{rtpp}, data from all sources are interpolated to a regular 50 m grid in the shifted height coordinate ($z - z_{\mathrm{rtpp}}$). The reference altitude, z_{rtpp}, does not affect the comparison between observation and model temperature and humidity on each profile. It only affects the composite obtained over many profiles due to the shift in each profile to the reference position, z_{rtpp}. The radiosonde-derived tropopause altitude is used because it is common to the comparisons with both the ECMWF IFS and MetUM models. This means that the sharp contrast between troposphere and stratosphere observed on most profiles is reflected in the composites and biases in the stratosphere can be clearly distinguished from those affecting the troposphere.

3.4 | Tropopause-matching condition

Figure 2 shows that on average the models and radiosondes generally agree in the altitude of the tropopause using the WMO lapse rate definition. The median tropopause altitude in the models is higher than that from the observed vertical profiles, by approximately 500 m for the MetUM at all lead times and approximately 200 m in

the IFS analysis increasing to 500 m in the five-day forecast. These differences are of a similar order of magnitude to the model grid spacing at these altitudes.

Figure 2 also shows that there are several cases where the diagnosed tropopause altitudes are very different between the model and the observations. The large differences in tropopause altitude between the observations and the analysis can be due to tropopause folds, with the lower tropopause being identified for one profile and the higher tropopause being identified for the other. In these situations, provided the model has represented the structure of the fold correctly, in agreement with observations, we would still expect any differences between modelled and observed profiles of temperature or humidity to represent errors in the model. The choice of z_{rtpp} used only affects the reference level used to composite the errors. Another cause of such large differences in tropopause height is feature displacement. This would occur if, for example, there was a difference in the forecast position of a Rossby wave on the tropopause at a given location such that the western side of a ridge was observed by the radiosonde, but the model profile was through the eastern side of the adjacent trough because the model had the wave slightly further east. In these situations we would also expect to see large differences in temperature and humidity for altitudes between those of the two different tropopause altitudes, as one profile would have stratospheric air here and the other tropospheric air. Such differences would not necessarily indicate an error in the model representation of the field, but rather displacement in the position of a large-scale feature, and so for the purposes of this study we will remove these cases. A cut-off is introduced such that, where the difference in tropopause height is greater than 1 km, the associated vertical profile is not included in these comparisons. This excludes those scenarios where the model and radiosonde profiles are through different sides of a sharp feature. This cut-off is only used in comparisons of observations to forecasts as it does not make a difference to the results of comparisons with the analysis, and including more cases allows us to produce better statistics. Data assimilation makes large feature displacement in the analysis unlikely. We can see from Figure 2 that the number of instances with large disagreement in tropopause altitude increases with forecast lead time, as feature displacement becomes more likely.

3.5 | Specific humidity normalised difference

The relative magnitude of the difference in specific humidity between model data and observations is shown using the unitless normalised difference between the model
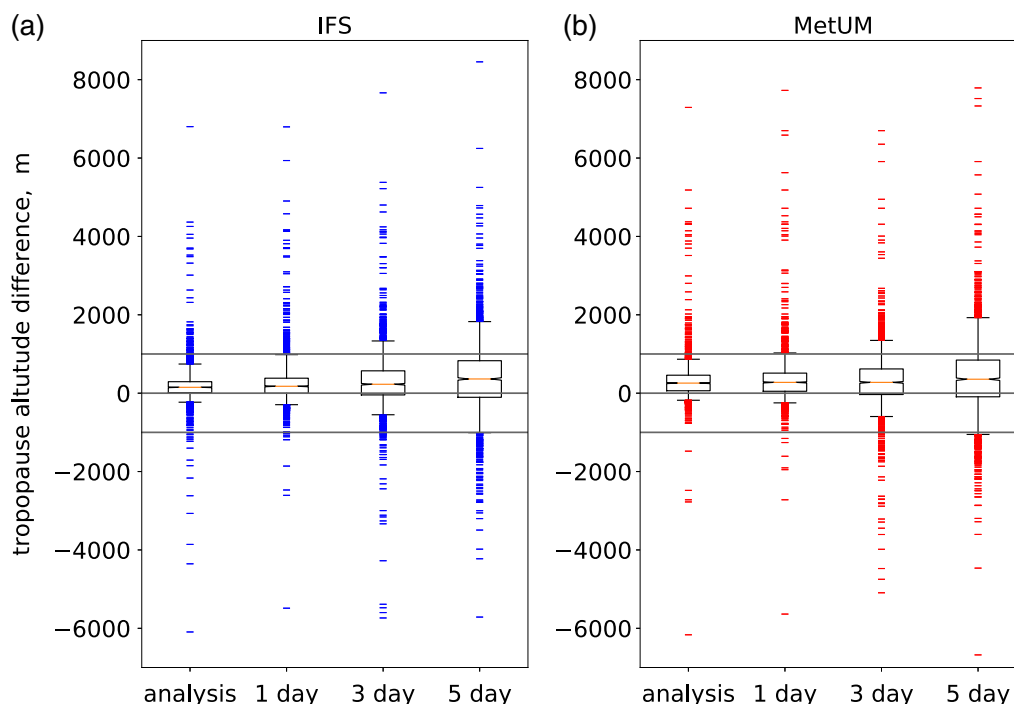
FIGURE 2 Box plots of the differences (model minus observed) in tropopause altitude as calculated from the model data and the radiosonde ascents for (a) IFS and (b) MetUM. Boxes in each panel are from left to right for analysis, 1-, 3- and 5-day forecasts. Boxes denote the interquartile range with the central red line indicating the median, the whiskers the 5th and 95th percentiles, and individual cases outside this range are shown with blue (IFS) and red (MetUM) markers. Horizontal grey lines denote differences of ±1 km [Colour figure can be viewed at wileyonlinelibrary.com]

humidity and the observed humidity, calculated as

$$\frac{q_{model} - q_{sonde}}{\sqrt{q_{model}^2 + q_{sonde}^2}}.$$

The advantages of normalising the differences by the root sum squared of the modelled and observed values are that it returns a value bounded between ±1 and is symmetric in magnitude with respect to relative differences of different sign between the model and radiosonde. Calculating the mean of this metric over a collection of ascents therefore does not give increased weight to those ascents with an overestimation of low humidity over those with an underestimation of higher humidity, as would occur if normalising by the observations alone. Some fractional differences given in the text are calculated from these mean normalised differences, for ease of interpretation.

# 4 | RESULTS OF COMPARING MODELS WITH RADIOSONDE OBSERVATIONS

In this section we first compare the radiosondes to the model analyses to find the magnitude and structure of

any systematic bias that exists in the initial conditions for forecasts, and investigate any dependence this has on the synoptic conditions. We then consider how these biases change over the first five days of the forecast.

4.1 | Comparison of observations to meteorological analyses

4.1.1 | Spatial and temporal variability and structure

We begin by examining the spatial and temporal structure of the differences between the radiosonde and IFS model data. Such comparisons with the MetUM yield similar results, and the mean biases of both models will be considered in later subsections. Figure 1 illustrates the positions of the radiosonde sites and gives a representative picture of how model temperature and humidity fields compare to observations on both sides of the tropopause in the UTLS. Figure 1a shows that the radiosonde observations of air temperature at 250 hPa agree with the background field from the IFS to within ±1.5°C, and there is a mixture of positive and negative differences, with no obvious
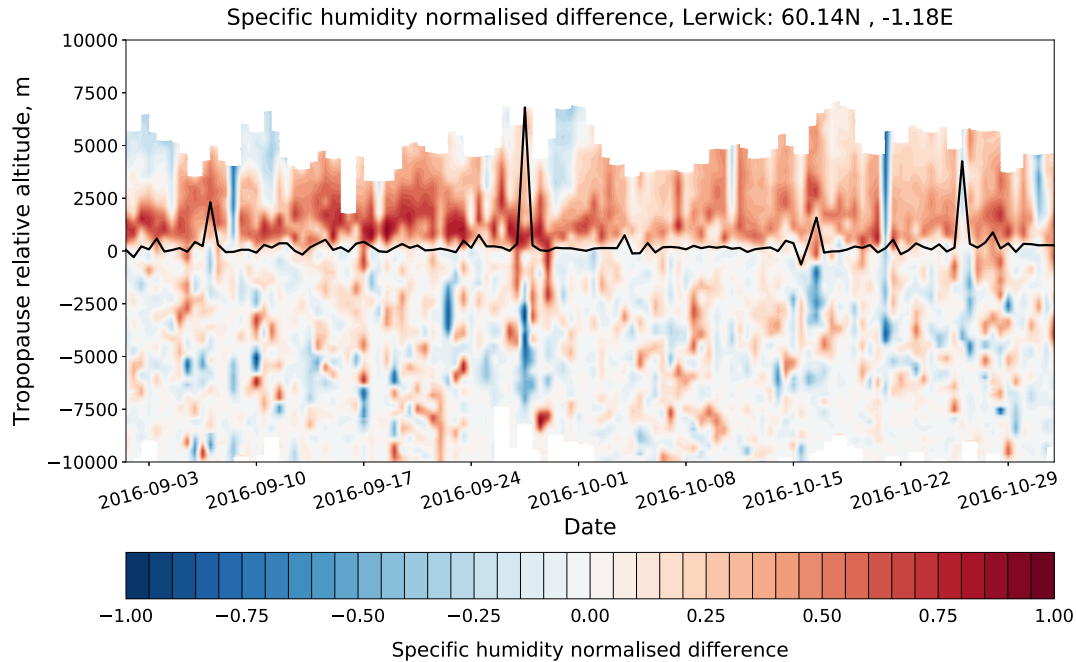
FIGURE 3 Timeseries over the two-month period from 01 September to 31 October 2016 at Lerwick (60.14°N, −1.18°E) of the normalised difference (colour shading) between IFS analysis specific humidity and radiosonde observations as a function of altitude. Altitude is considered in a tropopause-relative framework, where zero is the altitude of the tropopause as determined by the radiosonde observations, calculated according to the WMO lapse-rate definition as discussed in Section 3.3. The black contour is the tropopause as determined from the model data

suggestion of a systematic difference between analysed and observed temperatures. In contrast, in Figure 1b for specific humidity, the agreement is good in the troposphere, but the model has consistently higher humidity in the stratosphere.

The situation illustrated in Figure 1 of a moist bias in the lower stratospheric regions is representative of the entire two-month observation period considered. This persistence is shown in Figure 3 for a single observation site at Lerwick which used the RS92 radiosondes, though the results are consistent across the other sites considered. Below the tropopause there is variability in the specific humidity normalised difference, but no significant bias. In contrast, in the first few kilometres above the tropopause, the IFS has a systematic moist bias compared to the observations. The model tropopause is generally within a few hundred metres of that observed, with only a few outliers. For example, on 27 September the model tropopause is much higher than observed. Closer examination (not shown) reveals that this difference is due to a tropopause fold, as discussed in the previous section, and this sounding would be removed from the data used in the comparisons of the forecasts.

Considering now data from all locations and all times during the two-month period, in the scatter of observed against IFS model humidities (Figure 4) we see that in the troposphere for the majority of places and times the

humidities in both the model and the observations agree very closely (i.e., follow the 1:1 line). For the lowermost stratosphere, on the other hand, the model has a clear positive humidity bias compared to the observations. Taking those measurements made between 1 and 3 km above the tropopause and performing a linear regression on these, we recover a slope of 0.57, indicating that the IFS is 175% as moist as is observed, in the mean in this region. A further notable feature of Figure 4 is that values for the specific humidity in the model at altitudes above around 6 km above the tropopause seem to have a minimum at $\approx 3 \times 10^{-6}$. This is not as a result of an artificially imposed minimum value in the model, as we have found that the IFS is capable of sustaining lower values of specific humidity than this in forecasts (not shown).

4.1.2 | Mean vertical structure

We now consider the mean vertical structure of the model analysis humidity bias in an Eulerian frame of reference over the North Atlantic in the two-month period and in a tropopause-relative altitude coordinate system. For this analysis we will now also use data from the MetUM model in addition to that from the IFS. Additionally, as we are using observations made by radiosondes of two different types with different uncertainty characteristics,
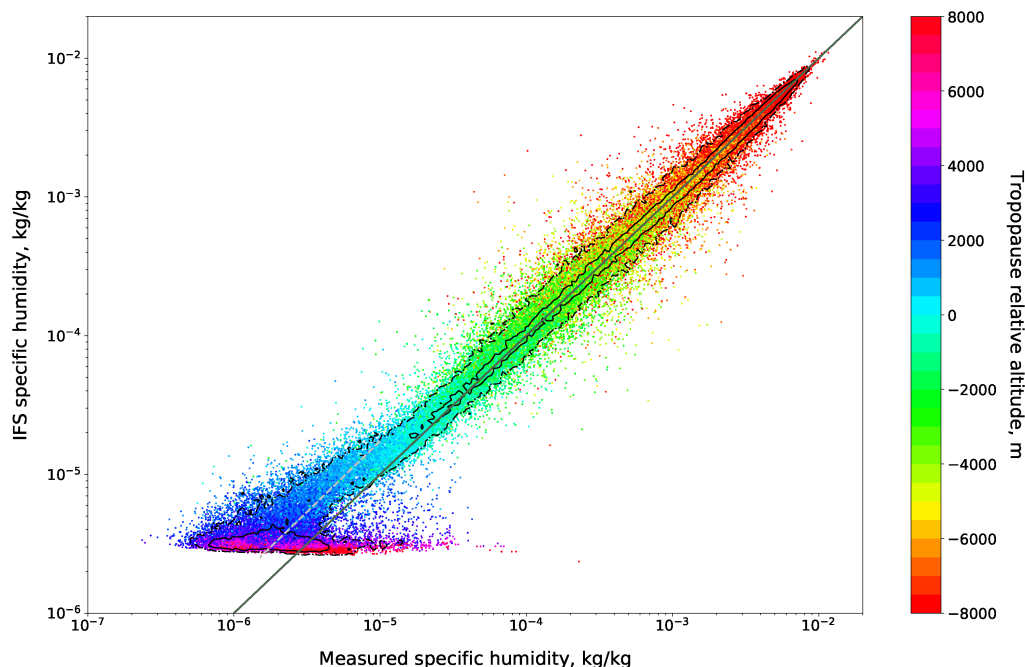
FIGURE 4 Average values of specific humidity in 500 m bins as measured by radiosondes (*x*-axis) against specific humidity at the same locations in IFS analyses (*y*-axis). Dots are coloured according to their tropopause-relative altitude. The solid grey line indicates a 1:1 relationship, and the dashed light grey line indicates the 0.57:1 relationship found considering a linear regression of points between 1 and 3 km above the tropopause. The black contours provide an indication of point density: points are five times more dense within the smaller solid contoured regions containing almost half of the points, than within the dot-dashed larger contour containing almost 90% of the points [Colour figure can be viewed at wileyonlinelibrary.com]

here comparisons of the models to these are composited separately. Figure 5a, c show that analyses from both models represent the specific humidity well in the troposphere, where observations are assimilated, but have mean moist biases in the stratosphere that increase in magnitude from small values at the tropopause to a maximum 1–2 km above the tropopause. There is also a sharp decrease across the tropopause in the observed profiles of specific humidity that is not replicated in the models. This decrease in the observations but not the models is still seen (though slightly less sharply) when compositing profiles relative to the model tropopause altitude (not shown) as opposed to the radiosonde tropopause altitude as shown here. This indicates that the difference between modelled and observed specific humidity across the tropopause is not caused by compositing model profiles with slightly different tropopause altitudes, but rather is a robust feature. These panels also give an indication of the data availability as a function of tropopause-relative altitude and show that this drops off with height in the stratosphere. In order to make a comparison relative to the tropopause altitude, all radiosonde profiles necessarily reach this level (ascents which do not cannot be used) and at some altitude above this the radiosonde balloons must burst.

Figure 5 illustrates how the measurement uncertainty of the radiosondes as given in Section 2.1 in terms of

relative humidity is related to uncertainty in specific humidity using the following method. For each ascent the relative humidity is calculated, upper and lower bounds on this are found according to the manufacturer-specified combined uncertainties quoted above, and these bounds are then converted back to specific humidity and averaged in the same way as specific humidity. As we are considering the mean over thousands of data points, the standard error of the mean is very small, assuming that systematic errors have been corrected for and the remaining measurement errors are random. The standard deviation shown here therefore does not indicate a lack of confidence in the magnitude of the bias, but rather illustrates the spread. Compared to the RS92 radiosondes in Figure 5b, the analyses at ≈1 km above the tropopause have a maximum mean normalised difference of ≈0.37. Compared to the RS41 radiosondes in Figure 5d, analyses have a maximum mean normalised difference of ≈0.34 at ≈1 km above the tropopause. A mean normalised difference of 0.34 means that the mean specific humidity in analyses is 166% of the mean observed value.

In the troposphere the RS41 radiosondes agree with model humidities, while observations from RS92 radiosondes are slightly drier. This is in agreement with the discussion of sensor differences in Section 2.1. Similarly we might expect measurements from RS92 radiosondes to
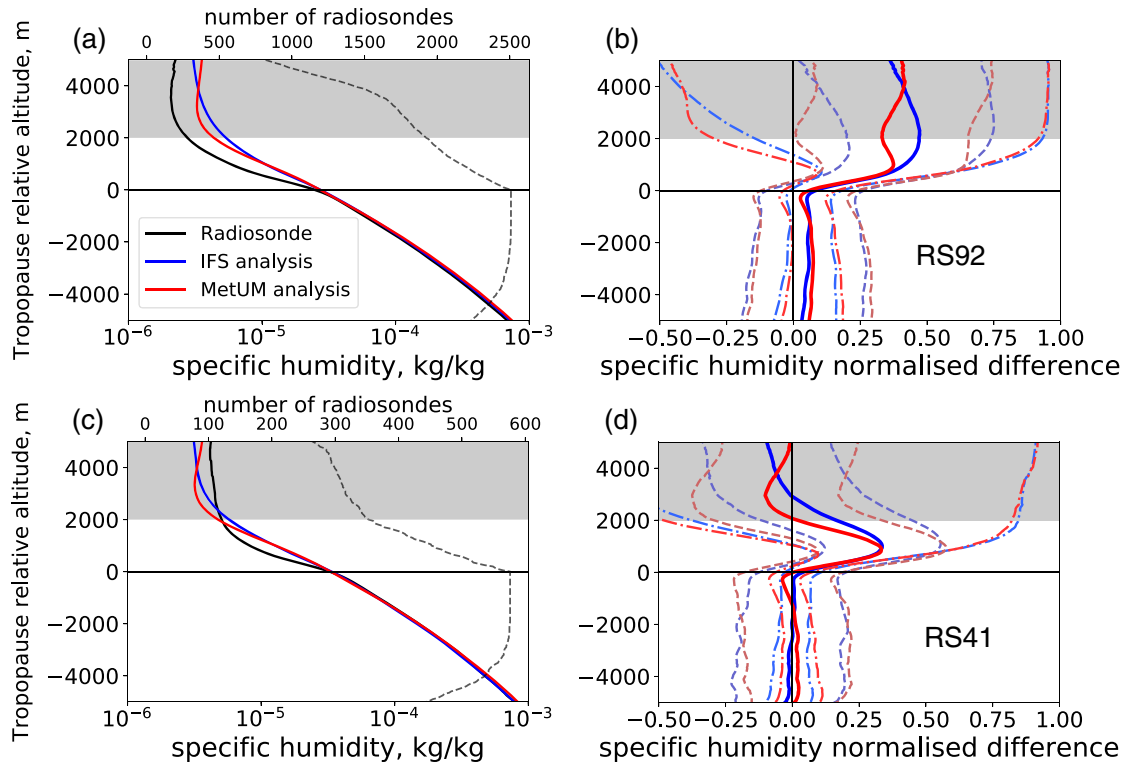
FIGURE 5 Mean tropopause-relative vertical profiles of (a, c) specific humidity and (b, d) the normalised difference of specific humidity, showing composites from sites using (a, b) the RS92 radiosonde and (c, d) the RS41 radiosonde. The black lines in (a, c) are for the radiosonde data, blue lines IFS analyses, and red lines MetUM analyses. The top *x*-axis is the scale for the number of points for which there are radiosonde data as a function of altitude, and hence the number of points contributing to the means at each tropopause-relative altitude, shown by the dashed grey line. In (b, d), dot-dashed lines show the measurement uncertainty as described in the text, and the dashed lines show one standard deviation of our sample of profiles of specific humidity normalised differences. Tropopause-relative altitudes above 2 km are grey shaded to indicate the large uncertainties in measurements of humidity here

be slightly moister than from RS41 in the lower stratosphere, resulting in a slightly reduced difference between the RS92 measurements and the model. In the lowermost stratosphere, the bias found through comparison to both radiosonde types is very similar, and even considering the uncertainty in the observations the lower limit of the normalised difference still indicates a moist bias 1 km above the tropopause in the analyses from both models. The character of the normalised specific humidity difference is different for the two radiosonde types higher up in the stratosphere. However, as is noted in Section 2.1, at altitudes above 2 km above the tropopause the instrument uncertainty is large. Furthermore, as the number of measurements at these altitudes is relatively small, and the RS41 and RS92 radiosondes are typically launched from different observing sites, from the available data we are unable to draw any conclusions regarding the comparison at these higher altitudes. Though values in this region are still plotted for consistency with the later consideration of temperature biases, for plots using humidity observations these regions are shaded grey. As the two radiosonde types perform similarly in our region of interest between

the tropopause and 2 km above it, data from the two radiosonde types are combined in the remainder of this paper. It should be noted that, as there are around five times more observations from RS92 than RS41 radiosondes, the former dominate subsequent statistics.

4.1.3 | Meteorological dependence of the vertical structure

To consider how the vertical structures and magnitudes of these biases vary under different atmospheric conditions, we can produce similar figures to Figure 5, but instead taking the mean only over profiles that satisfy certain criteria. This partitioning has been done conditioning on profiles that satisfy the following criteria: the presence of a low static stability layer within a certain distance below the tropopause; clouds within a certain distance of the tropopause; whether the vertical profile is taken through a ridge or a trough; and ridge profiles for which the air motion is northward (roughly equivalent to the western region of the ridge). The application of most of these
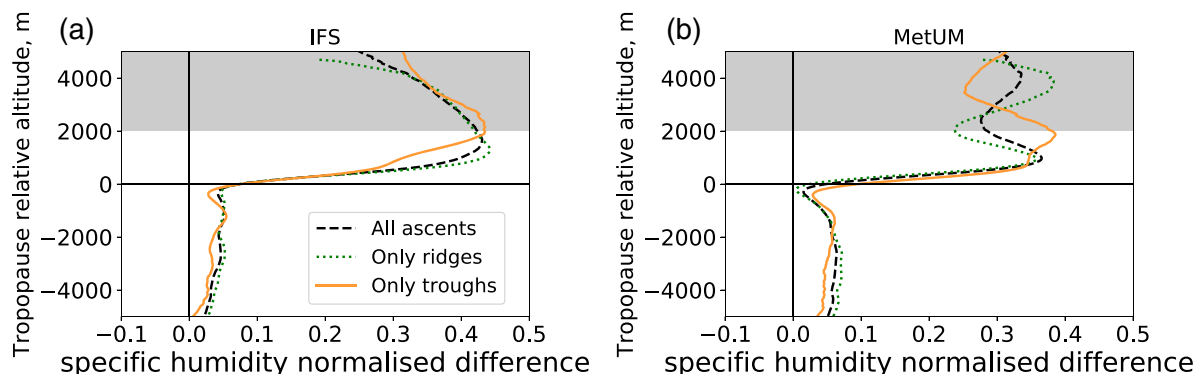
FIGURE 6 Mean tropopause-relative vertical profiles of the normalised difference of specific humidity for (a) IFS and (b) MetUM analyses. Black dashed lines show results over all ascents, green dotted only those in ridges, and orange solid only those in troughs. Grey shading is as in Figure 5 [Colour figure can be viewed at wileyonlinelibrary.com]

conditions resulted in no notable systematic difference between the composite vertical profiles of humidity (not shown) and are not discussed further here; the exception is when separating profiles in ridges and troughs. Ridges are identified relative to the mean height in ERA-Interim of the tropopause at each radiosonde site over the months of September and October from 2005–2017. The mean ERA-Interim tropopause height is calculated from the height of the 2 PVU surface which, due to the sharp PV gradient between the troposphere and stratosphere, is a convenient identifier of the tropopause in the Extratropics. A profile is classified to be in a ridge if the calculated tropopause altitude is larger than one standard deviation above this mean height. The standard deviation used is that of the set of 26 monthly means (from two months in each of 13 years). Troughs are identified similarly, but using one standard deviation below this mean height. Of the 3,204 vertical profiles, 1,605 of these were classified as ridges, and 772 as troughs.

The humidity biases have similar vertical structures in ridges and troughs, but with a larger vertical length-scale for troughs, and smaller for ridges (Figure 6). This results in the maximum difference being found at ≈1 km above the tropopause in ridges, but ≈2 km above in troughs. Additionally, although in the IFS the moist biases in the troughs and ridges are of very similar magnitude, there is a more pronounced difference in the MetUM, with the troughs having a slightly larger and much deeper systematic bias.

## 4.2 | Comparison of forecasts to observations

Having considered the comparison of these radiosonde observations to model analyses, we now also compare to forecasts out to five days. The specific humidity bias is

large in the analyses, and we can see from Figures 7a, b that the mean analysis value of the normalised difference in the lowermost stratosphere is much larger than any changes that occur in the first five days of the forecasts. In the IFS the mean moist bias 1 km above the tropopause increases slightly over the forecast period, whereas in the MetUM it decreases slightly.

Additionally we see that, although the temperature in the model analyses is very similar to that from the radiosonde observations, it diverges from the observations during the first five days of a forecast in both the IFS and the MetUM. Initially in both models there is a slight cold bias in the lowermost stratosphere of ≈0.2 K with respect to the radiosondes, and a slight warm bias at the tropopause of similar magnitude. Subsequently, the difference from observations of the lowermost stratosphere temperature evolves differently in the IFS and MetUM (compare Figure 7c, d). For the IFS a cold bias develops in the lowermost stratosphere at a rate of ≈0.2 K·day^{-1}; the largest bias is initially 600 m above the tropopause, but increases to 1,200 m above the tropopause during the five-day forecast. In the MetUM, the slight mean cold bias at ≈1 km height grows at a slower rate than for the IFS, and a mean warm bias develops at an altitude of ≈3 km above the tropopause; this bias develops at an accelerating rate, reaching a magnitude of ≈0.7 K after five days. Another feature in Figure 7c, d is a peak in the air temperature difference at the tropopause. This sharp peak at the tropopause is not present when compositing differences relative to the model tropopause (not shown). If tropopause altitudes between model and observations differ, the consequence of compositing model temperature profiles relative to the observed tropopause is that the transition from tropospheric to stratospheric lapse rates at the tropopause will be smoothed in the vertical, and the temperature in the narrow region about the tropopause in the model will be slightly increased. However there
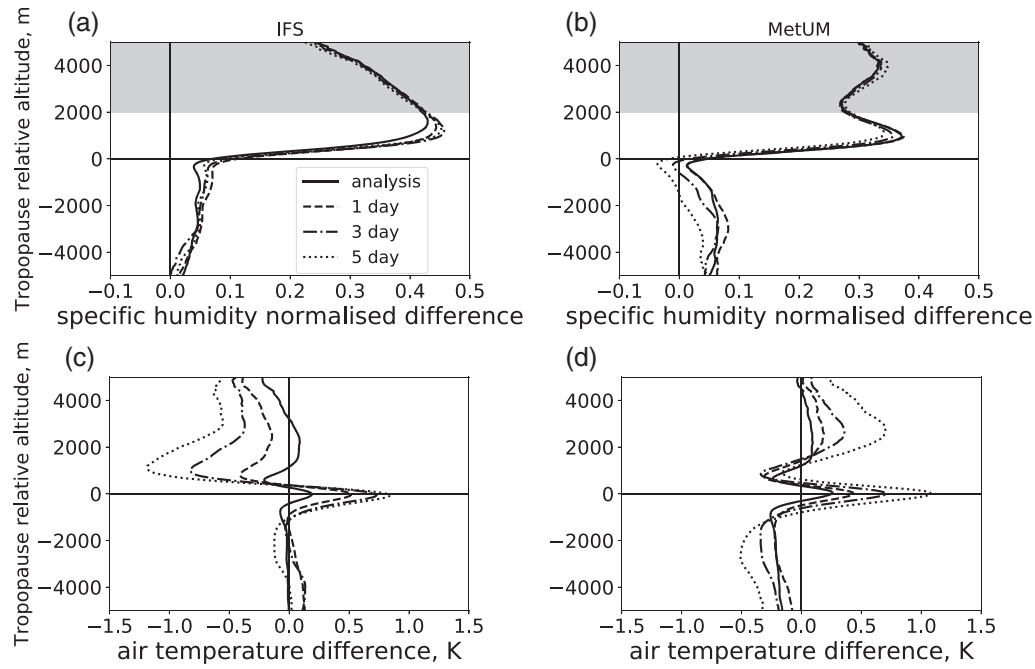
FIGURE 7 Mean tropopause-relative vertical profiles of (a, b) specific humidity normalised difference and (c, d) air temperature difference for (a, c) IFS and (b, d) MetUM. Solid lines denote differences with model analyses and the other lines differences with forecasts: dashed lines 1-day, dot-dashed lines 3-day, and dotted lines 5-day forecasts. Grey shading is as in Figure 5
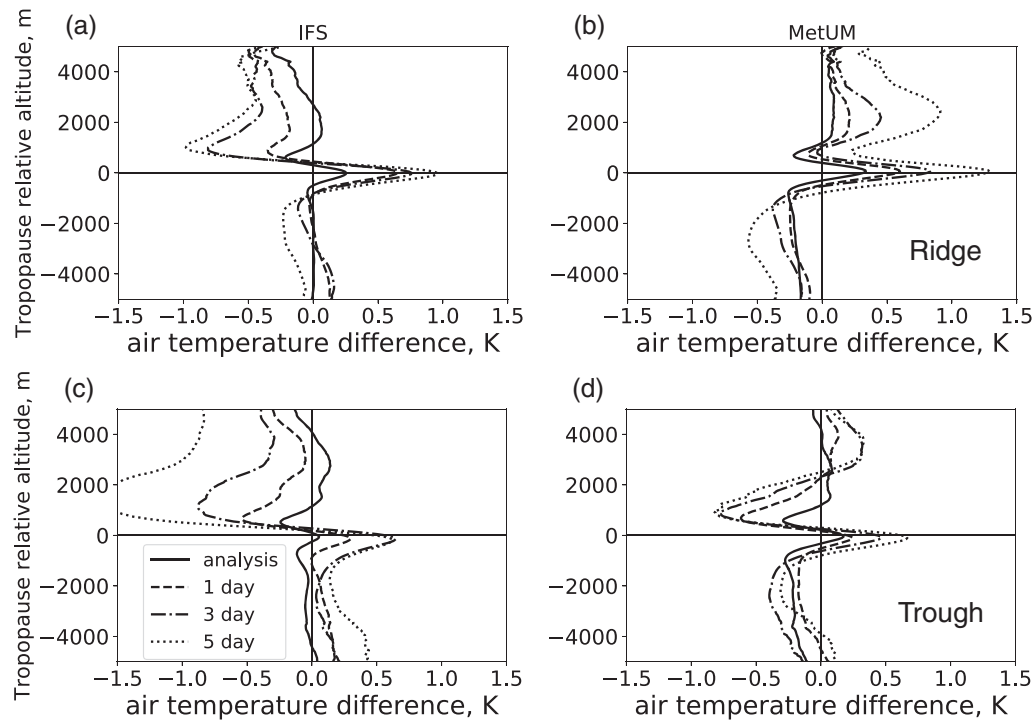
FIGURE 8 As Figure 7(b, d): mean tropopause-relative vertical profiles of air temperature difference for (a, c) IFS and (b, d) MetUM. (a, b) show means only over those ascents in ridges and (c, d) only those in troughs

is instead in a model–tropopause relative comparison a slight warm bias developing at a rate of approximately $0.1\,\text{K·day}^{-1}$ in a much deeper region below the model tropopause.

It is shown in Saffin *et al.* (2017) that the behaviours of physical processes affecting PV at the tropopause are different in ridges and troughs. Separating the air temperature difference profiles from Figure 7c, d into

ridges and troughs yields Figure 8. The greater number of profiles through ridges compared to troughs (more than twice as many) contribute to the mean profiles in Figure 7c, d bearing a closer resemblance to those shown for ridges than troughs. For the IFS, the magnitude of the peak temperature difference is very similar in ridges and troughs in the analyses and one- and three-day forecasts (≈ 0.85 K at 1 km above the tropopause). Recall that the specific humidity normalised difference was similar between ridges and troughs for the IFS (Figure 6a). However, whereas the maximum value of the normalised difference in specific humidity was at a higher tropopause-relative altitude for troughs than ridges, the altitude of the maximum cold bias is the same in both ridges and troughs, but the cold bias extends further up into the stratosphere in troughs.

In the MetUM we see larger differences when separating profiles into those in ridges and troughs, with a cold bias of around 0.8 K at 1 km above the tropopause increasing over the first three days of the forecast in troughs, as with the IFS, but a warm bias developing with lead time above ridges at 2–3 km above the tropopause.

Now we will consider briefly the sensitivity of these results to the choice of tropopause-matching condition (not shown). Results for humidity are not sensitive to the threshold chosen, nor are results for IFS temperature in the analysis and one- and three-day forecasts. However there is an increase in the cold bias from ≈ 0.2 to ≈ 0.3 K·day^{-1} in IFS five-day forecasts with the inclusion of cases where the model tropopause is between 1 and 1.5 km above that observed. MetUM temperature forecasts are also sensitive to the choice of tropopause matching condition. Regardless of the chosen threshold, we see the development of a warm bias 3 km above the tropopause in ridges and a cold bias 1 km above the tropopause in troughs, however in troughs without the condition the mean cold bias is markedly larger at both 3 and 5 days, and in the mean over all profiles when the matching condition is not applied there is a reduced warm bias at 3 km, and there is instead a larger cold bias at 1 km. It is for this reason that we are able to find a systematic tropopause-relative warm bias in the lowermost stratosphere of the MetUM when using this tropopause matching condition, while in climatological comparisons it is known that the MetUM has a cold bias in the extratropical lower stratosphere (Hardiman *et al.*, 2015). It is shown in Figure 2 that there are more cases where the model tropopause is higher than observed excluded by the matching condition than lower, increasing with lead time, and these differences when removing the tropopause matching condition are consistent with this. The relation between the biases and tropopause altitude will be discussed further in Section 5.2.

5 | ATTRIBUTING TEMPERATURE BIASES TO THE HUMIDITY BIAS

In the Introduction we discussed the mechanism by which a moist bias in the lowermost stratosphere can lead to a collocated cold bias through changes in long-wave radiative transfer. In Section 4, we identified the locations and magnitudes of both a moist bias in the model analyses and the development of a temperature bias in forecasts. In this section we examine what proportion of the forecast temperature biases can be attributed to changes in radiative cooling, as a consequence of the presence of the moist bias in the analyses, and consequent changes in the structure of the tropopause.

5.1 | Attribution of temperature biases using the IFS single-column model

The magnitude of the cold bias attributable to the radiative response to the identified moist bias is quantified using the IFS SCM. The SCM simulations are used to provide information on the typical heating rates from parametrized physical processes, the contributions of which are output as individual rates of change of temperature, referred to as "process tendencies". Although in the full 3D forecast models air masses are advected around, such that over a five-day forecast they will experience heating or cooling in different regions, the SCM just represents the impact of the physical processes in a column with no advection. Two simulations are run for nine days initialised with identical temperature profiles, but humidity profiles representative of the IFS analysis state and the observed state respectively. The differences in temperature between these are therefore as a result of the differing initial humidity profiles. The SCM simulation representative of the IFS model state (labeled IFS-prof in Figure 9) is initialised with a humidity profile equivalent to the mean over the humidity profiles from the IFS analyses at all radiosonde launch sites and times. This profile is similar to that shown in Figure 5a,c, but the mean is taken in a ground-relative sense in pressure coordinates, as opposed to a tropopause-relative sense in altitude coordinates. The SCM is also initialised with a similarly produced temperature profile, zero wind speed, and a surface pressure of one atmosphere: 101,325 Pa. The surface boundary condition is zero sensible and latent heat flux.

The simulation representative of the atmosphere as observed by the radiosondes (labelled Obs-prof in Figure 9) is initialised with the same temperature as those used for the IFS simulation, but with the humidity profile changed only in the lowermost stratosphere,
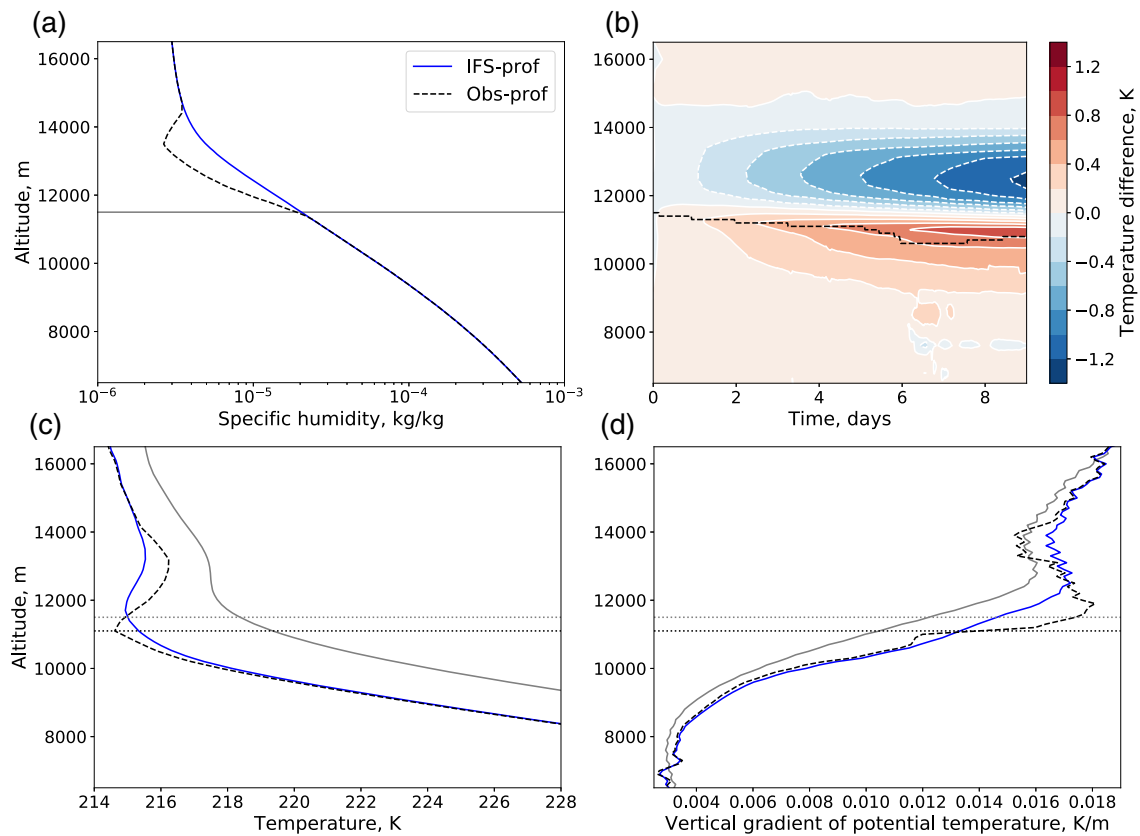
FIGURE 9 (a) Idealised vertical profiles of specific humidity representative of the IFS (blue) and radiosonde observations (black dashed) used for initial conditions of the SCM simulations, as described in the text. The horizontal grey line is the tropopause as calculated from the mean temperature profile using the WMO definition. (b) The temperature in the simulation with IFS humidity minus that with the humidity modified towards that observed, for the nine days of the forecast. The black dotted line is the tropopause as calculated from the temperature profiles of the Obs-prof simulation and dashed contours indicate negative values. Vertical profiles of (c) temperature and (d) the vertical gradient of potential temperature at five days for IFS (blue, IFS-prof) and radiosonde (black dashed, Obs-prof) representative SCM simulations, and the profile at initial time (grey) [Colour figure can be viewed at wileyonlinelibrary.com]

as indicated in Figure 9a. In Section 4 we identify the mean tropopause-relative specific humidity normalised difference over all considered radiosonde ascents, illustrated by the solid line in Figure 7a. The observation-representative humidity profile is created such that the tropopause-relative normalised difference between Obs-prof and IFS-prof between the tropopause and 2 km above is the same as that difference found in Section 4. This is achieved by removing the appropriate amount from the IFS-prof averaged humidity profile. This layer is chosen to isolate the effect of a moist bias in the lowermost stratosphere where we have the highest confidence in the presence of a moist bias. The number of radiosonde measurements above this level is small and the measurement uncertainty is much greater, so a humidity bias cannot be robustly confirmed above this level. Below the tropopause and above 3 km above the tropopause, the imposed normalised difference profile between Obs-prof and IFS-prof is set to be zero, and it varies linearly between the mean value at 2 km and zero at 3 km. A linear

reduction of the normalised difference in this layer is imposed to prevent any effects arising from a sharp boundary. Such a layer is not needed below the tropopause as the normalised difference here is already approximately zero. Results from SCM simulations initialised with these idealised mean profiles are representative of results from SCM simulations initialised using profiles from individual times and locations.

Figure 9(b) illustrates the difference in temperature fields between the two SCM simulations. From examination of the temperature tendencies (not shown), we know that this difference is entirely attributable to the radiation scheme. A dipole of temperature difference emerges over the nine days of forecasts shown, with a lowermost stratosphere cold bias and near-tropopause warm bias. The additional cooling of IFS-prof relative to Obs-prof is -0.175 K·day^{-1} at day 3, reducing to -0.16 K·day^{-1} at day 5 and continuing to slow for the remaining four days of the simulation. If an alternative scenario is assumed where the moist bias extends upward and humidity is reduced to

match observations in the lowest 5 km of the stratosphere and linearly to no difference at 7.5 km (not shown), then the cooling rate slightly decreases to $\approx -0.14\,\text{K·day}^{-1}$ at day 5. The warm bias in the vicinity of the tropopause, and just below, is shallower and of a lower magnitude than the cold bias above, growing at a similar rate of $\approx +0.15\,\text{K·day}^{-1}$.

The increased long-wave cooling in the lowermost stratosphere associated with the moist bias leads to the growth of the negative temperature bias in these forecasts locally. However, the increased down-welling long-wave emission from the lowermost stratosphere will be absorbed in the upper troposphere (Birner and Charlesworth, 2017), leading to increased warming in the upper troposphere with respect to a profile without a moist bias. Consequently, we expect the upper tropospheric warm bias to grow in tandem with the lower stratospheric cold bias, as seen in the SCM (Figure 9).

The SCM-derived cooling rate above the tropopause is consistent with the $-0.2\,\text{K·day}^{-1}$ found in the mean through comparison of operational forecasts to radiosonde temperature observations in Figure 7, given the idealised nature of the SCM simulations. This consistency strongly suggests that the cold bias which is found to develop in the lowermost stratosphere of the IFS is largely a result of the moist bias in the lowermost stratosphere analysis.

The growth of a warm bias is also seen in both operational models (Figure 7) but is more concentrated at the tropopause than in the SCM. It is also found to be less prominent if the model–observation differences are composited relative to the tropopause in the forecasts, rather than using the observations as the reference (not shown). This difference in the temperature bias between the SCM and operational model suggests there are other processes acting in the full IFS forecasts, such as advection and mixing, which modify the warm bias in the region of the tropopause.

5.2 | Changes in static stability near the tropopause

Water vapour in the lowermost stratosphere is also shown to affect tropopause altitude and sharpness. The temperature profile of the Obs-prof SCM simulation, initialised with a drier lowermost stratosphere, evolves to have a sharp transition from positive to negative lapse rates at the tropopause, above which the temperature then increases to a local maximum ≈ 2 km above the tropopause at the top of the TIL, before becoming approximately isothermal for several kilometres (Figure 9c). This structure is similar to that of temperature profiles from the observations
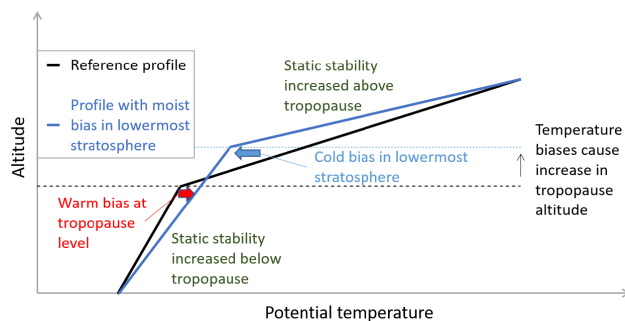
FIGURE 10 Schematic illustration of the effect of the lowermost stratospheric moist bias on an idealised vertical profile of potential temperature, where the black line is an idealised reference profile. The blue line represents the profile following the effects of the anomalous radiative heating dipole that results from the presence of a moist bias. The horizontal dotted lines indicate the tropopause altitudes for the respective profiles [Colour figure can be viewed at wileyonlinelibrary.com]

and analysis (not shown). However, with the moist bias in IFS-prof there is a weaker lapse rate in the upper troposphere, a smoother transition from positive to negative lapse rates, the local minimum in temperature is at a higher altitude, and the magnitude of the local temperature maximum at the top of the TIL is reduced. This, on the other hand, is similar to that of temperature profiles from the operational five-day forecasts. That we see these similarities, despite the SCM being initialised with a smoothed temperature profile, demonstrates the dependence of the structure of the equilibrium model temperature profile at the tropopause on initial lowermost stratospheric humidity.

The schematic in Figure 10 illustrates how the radiative heating dipole in response to the moist bias is expected to influence an idealised atmospheric profile where static stability is piecewise uniform. To first order, the tropopause can be represented as a sharp change in static stability from typical values in the troposphere to a higher value in the lower stratosphere (approximately double). The air cools radiatively where there is the moist bias and potential temperature increases below (near the tropopause). Local turbulent mixing tends to maintain uniform static stability in the troposphere and the blue curve in Figure 10 is obtained by satisfying the radiative changes in potential temperature and piecewise uniform static stability. A necessary result is an increase in tropopause height, as well as increased static stability above the level of peak anomalous cooling, as seen in Figure 9d. Static stability must also decrease in the model between the level of the observed tropopause and the peak of anomalous cooling. This tropopause smoothing is illustrated by the SCM results in Figure 9 at the level of the Obs-prof tropopause,

showing that, in models with a lowermost stratosphere moist bias in the initial conditions, there is a reduction of the sharp static stability gradient.

This can explain plausibly why the tropopause altitude calculated from model data is on average higher than in radiosonde observations, and that this difference increases with forecast lead time (Figure 2a).

5.3 | Unattributed temperature biases

Finally, the temperature biases that developed in the MetUM forecasts are considered.

In contrast to the findings for the IFS operational forecasts, the SCM experiments indicate that the radiative response to a moist bias cannot completely explain the temperature biases in the MetUM operational forecasts. Additional SCM simulations were run using idealised humidity profiles representative of the mean over those from the MetUM as well as taken from individual ascents. Through comparisons of temperature fields from these simulations (not shown) it is evident that the a lowermost stratosphere cold bias develops due to the effects of long-wave radiation in a similar way to the IFS, although with a smaller magnitude. In Figures 7d and 8b,d the temperature difference 1 km above the tropopause is more negative than that 3 km above the tropopause, which we would expect from radiative effects of the moist bias. The warm bias 3 km above the tropopause in ridges, in conjunction with the smaller negative temperature bias below this, may then be the result of a secondary factor causing a warming throughout the stratosphere in ridges. It is hypothesised that another of the dominant stratospheric trace gases (i.e., ozone or carbon dioxide) is responsible for this additional anomalous warming. Both the IFS and the MetUM use an ozone climatology which does not take into account the varying tropopause height. As a result of this, both have higher concentrations of ozone in the lower stratosphere above ridges than observed at the same altitudes by ozonesondes (World Ozone and Ultraviolet Radiation Data Centre, 2020) or the AIRS satellite instrument (Teixeira and AIRS Science Team, 2013). At most latitudes this difference is notably larger than the difference between the climatologies used by the two models. So, although one might expect a positive bias in lower stratospheric ozone concentrations in ridges in the MetUM to cause anomalous warming there, it is unclear why this would would result in a warm bias in the MetUM (at 3 km above the tropopause), but not in the IFS. As mentioned previously, this temperature bias will be affected by the history of the air parcels as well as local radiative processes.

6 | SUMMARY AND CONCLUSIONS

An accurate representation of temperature and humidity in the extratropical lower stratosphere is important for global weather forecast and climate models. Many models have been shown to have significant biases in this region and there has been little progress in reducing these biases. To improve the models, it is imperative to gain a better understanding of the errors and their sources and this study is a step forward in that direction. The specific aims of this study are to identify and characterise humidity and temperature biases in the upper troposphere and lowermost stratosphere in the IFS and MetUM weather forecast models using radiosonde observations, to quantify the temperature bias growth attributable to a diagnosed moist bias, and to explore the influences of these biases on other tropopause-level features. The main conclusions are summarised below.

It is found that both the IFS and MetUM have a mean moist bias in the lowermost stratosphere with a maximum approximately 1 km above the tropopause with the model humidities approximately 170% of the observed values. The magnitude of this bias found through comparison with radiosonde data is largely consistent with those found by the previous studies listed in Table 1 derived from comparison with aircraft and satellite observations, and the altitude is consistent with the autumn tropopause-relative comparison from Dyroff *et al.* (2015). When considering only radiosonde ascents through tropopause-level troughs, this maximum value of the specific humidity normalised difference occurs around 2 km above the tropopause; for ridges it occurs at around 1 km, the same height as in the mean. In the IFS, the tropopause-relative vertical structure of the moist bias is very similar between ridges and troughs, whereas in the MetUM features in troughs have a larger vertical length-scale. The moist bias is not found to be systematically dependant on the presence of cloud, layers of low static stability, or position within a ridge, and the magnitude of the moist bias changes very little during five-day forecasts in both models, being dominated instead by the bias present initially.

The temperature fields in model analyses agree with radiosonde observations to within ± 0.2 K. However, a cold bias develops in the lowermost stratosphere in the IFS operational forecasts, also with a maximum approximately 1 km above the tropopause and growing at a rate of ≈ 0.2 K$\cdot$day^{-1}. There is little difference in this growth rate between troughs and ridges in the first three days, although the cold bias in troughs extends deeper into the lowermost stratosphere. Using the IFS single-column model (SCM) it is shown that the growth of this lowermost stratosphere cold bias is consistent with the additional long-wave radiative cooling calculated as a result of the

presence of the lowermost stratosphere moist bias. The SCM simulations also show that the moist bias would result in a warming around the tropopause level, extending below into the upper troposphere, with a smaller depth and growth rate than the lowermost stratospheric cooling. A warming is also seen in the operational forecasts as a sharp feature at the tropopause, but this does not extend into the upper troposphere in the same way as the warming in the SCM. This is likely to be due to advection and mixing processes which are not represented in the SCM but modify the warming feature in the 3D model. However this feature in the composites is also sensitive to the specification of tropopause-relative coordinates.

As is the case for the IFS, the MetUM operational forecasts also develop a cold bias at 1 km above the tropopause, but with a smaller magnitude than that in the IFS and not present when considering only profiles in ridges. For ridges instead a warm bias develops with a maximum at around 2.5 km above the tropopause. This warm bias is seen too in the mean over all profiles, corresponding to an additional warming of ≈ 0.1 K·day^{-1}. We cannot explain the MetUM lower stratosphere warm bias as a radiative response to the moisture bias. Indeed, additional long-wave cooling would be expected due to the moist bias in both models. Although the static ozone climatology would be expected to generate a warm bias above large-scale tropopause ridges, this does not explain why the responses of the MetUM and IFS differ. Therefore, further investigation would be required to understand the development of the warm bias in the full MetUM model. In addition to temperature biases, it is demonstrated that the presence of a lowermost stratosphere moist bias results in a higher tropopause, a less pronounced TIL, and a smoother cross-tropopause static stability gradient than in an atmospheric profile with a drier lowermost stratosphere (summarised in Figure 10). As discussed in the Introduction, such changes to the structure of the tropopause would be expected to systematically affect wave propagation on this sharp gradient, and through this other aspects of forecast development in the troposphere.

A limitation of the results regarding the magnitude and structure of the found biases is that data were only considered over the eastern North Atlantic and western Europe region, in a two-month period. These data were chosen to more easily facilitate future studies of the particular impact of these biases on the development of extratropical cyclones in this region, and consequently on forecast quality. Due to the NAWDEX campaign, the high-resolution radiosonde data were readily available. However, as is shown by Dyroff *et al.* (2015), the lowermost stratosphere moist bias in the IFS varies seasonally with a maximum in summer and minimum in winter. To further understand the sources of error and their seasonal variations,

temperature and humidity biases could be similarly analysed in other seasons. Furthermore, although the temperature data are reliable at all considered altitudes, due to the measurement uncertainty in the instruments we have confidence in our assessment of the moist bias only in the lowest 2 km of the stratosphere. To better understand model performance in humidity representation at higher altitudes than these, it would be valuable for complementary studies using other observation techniques to be performed.

Previous work has shown that a moist bias in the extratropical lowermost stratosphere would be expected to cause a collocated cold bias (Forster and Shine, 2002; Maycock *et al.*, 2011). This study has found that the growth rate of this lowermost stratosphere cold bias in the ECMWF operational forecast model is quantitatively consistent with the additional radiative cooling rate calculated using the magnitude of the moist bias found in the IFS analysis. We have shown that the moist bias is dominated by the bias in the initial conditions and therefore, to reduce this cold bias in forecasts, reduction of the moist bias in the analysis is required. However, as increments are not currently applied to humidity fields in this region during data assimilation, as discussed in Section 2, the analysis humidity field is not constrained by observations. Whether or not humidity observations are assimilated, it is important to understand and reduce the sources of the moist bias in the forecast model. It is likely that there are contributions from excessive diffusion or transport of water vapour across the hygropause from the high water vapour values in the troposphere due to errors in numerical or physical processes. The characterisation of the magnitude and vertical structure of the model moist bias presented in this paper facilitates a more detailed follow-on investigation of model processes controlling humidity in this region which would be required to address this problem.

AUTHOR CONTRIBUTIONS

Jake Bland: conceptualization; formal analysis; investigation; writing – original draft; writing – review and editing. **Suzanne Gray:** conceptualization; funding acquisition; supervision; writing – review and editing. **John Methven:** conceptualization; funding acquisition; supervision; writing – review and editing. **Richard Forbes:** conceptualization; resources; supervision; writing – review and editing.

ACKNOWLEDGEMENTS

additional launches during the NAWDEX campaign; we also thank the principal investigator of the EUMETNET project for NAWDEX, Andreas Schäfler (Deutsches Zentrum für Luft- und Raumfahrt). The authors also thank the ECMWF for data access, and Claudio Sanchez for obtaining the Met Office analyses and forecasts which were used here. We are very grateful to Ben Harvey for his help with the radiosonde and MetUM data, and for discussion on smoothing, low static stability layers and tropopause folds. Further thanks to Keith Shine for useful discussion of radiation. We are also very grateful to Bruce Ingleby and Roger Saunders for correspondence regarding the assimilation of radiosonde data at the ECMWF and Met Office. We additionally thank Maria Broadbridge for invaluable help in using the SCM. Thanks also to two anonymous reviewers for their comments. J.B. was funded by the Natural Environment Research Council (NERC) SCENARIO Doctoral Training Partnership (NE/L002566/1).

ORCID
Jake Bland https://orcid.org/0000-0003-2706-2853
Suzanne Gray https://orcid.org/0000-0001-8658-362X
John Methven https://orcid.org/0000-0002-7636-6872
Richard Forbes https://orcid.org/0000-0002-3596-8287

REFERENCES

Bannister, R.N., Chipilski, H.G. and Martinez-Alvarado, O. (2020) Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales. *Quarterly Journal of the Royal Meteorological Society*, 146, 1–48.

Birner, T. (2006) Fine-scale structure of the extratropical tropopause region. *Journal of Geophysical Research: Atmospheres*, 111(D4). https://doi.org/10.1029/2005JD006301.

Birner, T. and Charlesworth, E.J. (2017) On the relative importance of radiative and dynamical heating for tropical tropopause temperatures. *Journal of Geophysical Research: Atmospheres*, 122(13), 6782–6797.

Carminati, F., Migliorini, S., Ingleby, N.B., Bell, W., Lawrence, H., Newman, S., Hocking, J. and Smith, A. (2019) Using reference radiosondes to characterise NWP model uncertainty for improved satellite calibration and validation. *Atmospheric Measurement Techniques*, 12(1), 83–106.

Carver, G. (2019). Single-column model 43r3 release notes; OpenIFS – ECMWF Confluence Wiki. https://confluence.ecmwf. int/display/OIFS/Single+column+model+43r3+release+notes; accessed 26 April 2021.

Chagnon, J.M., Gray, S.L. and Methven, J. (2013) Diabatic processes modifying potential vorticity in a North Atlantic cyclone. *Quarterly Journal of the Royal Meteorological Society*, 139, 1270–1282.

Dee, D.P. and Da Silva, A.M. (2003) The choice of variable for atmospheric moisture analysis. *Monthly Weather Review*, 131(1), 155–171.

Diamantakis, M. and Flemming, J. (2014) Global mass fixer algorithms for conservative tracer transport in the ECMWF model. *Geoscientific Model Development*, 7, 965–979.

Dirksen, R.J., Bodeker, G.E., Thorne, P.W., Merlone, A., Reale, T., Wang, J., Hurst, D.F., Demoz, B.B., Gardiner, T.D., Ingleby, N.B., Sommer, M., Von Rohden, C. and Leblanc, T. (2020) Managing the transition from Vaisala RS92 to RS41 radiosondes within the GCOS Reference Upper-Air Network (GRUAN): a progress report. *Geoscientific Instrumentation, Methods and Data Systems*, 9, 337–355. https://doi.org/10.5194/gi-9-337-2020.

Dyroff, C., Zahn, A., Christner, E., Forbes, R., Tompkins, A.M. and van Velthoven, P.F.J. (2015) Comparison of ECMWF analysis and forecast humidity data with CARIBIC upper troposphere and lower stratosphere observations. *Quarterly Journal of the Royal Meteorological Society*, 141, 833–844.

ECMWF (2016). IFS documentation CY41r2, ECMWF, Reading, UK. Available at: https://www.ecmwf.int/en/publications/ifs-documentation.

ECMWF (2017). IFS Documentation CY43R3, Part IV: physical processes, ECMWF, Reading, UK.

Edwards, D., Anderson, G., Oakley, T. and Gault, P. (2014). Met Office intercomparison of Vaisala RS92 and RS41 radiosondes. Technical report, Exeter, UK: Met Office.

Feist, D.G., Geer, A.J., Müller, S. and Kämpfer, N. (2007) Middle atmosphere water vapour and dynamical features in aircraft measurements and ECMWF analyses. *Atmospheric Chemistry and Physics*, 7, 5291–5307.

Forster, C. and Wirth, V. (2000) Radiative decay of idealized stratospheric filaments in the troposphere. *Journal of Geophysical Research: Atmospheres*, 105(D8), 10169–10184.

Forster, P.M.F. and Shine, K.P. (1999) Stratospheric water vapour changes as a possible contributor to observed stratospheric cooling. *Geophysical Research Letters*, 26(21), 3309–3312.

Forster, P.M.F. and Shine, K.P. (2002) Assessing the climate impact of trends in stratospheric water vapor. *Geophysical Research Letters*, 29(6), 10–1–10-4.

Fowler, A., Bannister, R.N. and Eyre, J.R. (2012) A new floating model level scheme for the assimilation of boundary-layer top inversions: the univariate assimilation of temperature. *Quarterly Journal of the Royal Meteorological Society*, 138, 682–698.

Gates, W.L., Boyle, J.S., Covey, C., Dease, C.G., Doutriaux, C.M., Drach, R.S., Fiorino, M., Gleckler, P.J., Hnilo, J.J., Marlais, S.M., Phillips, T.J., Potter, G.L., Santer, B.D., Sperber, K.R., Taylor, K.E. and Williams, D.N. (1999) An overview of the results of the Atmospheric Model Intercomparison Project (AMIP I). *Bulletin of the American Meteorological Society*, 80, 29–56.

Gray, S.L., Dunning, C.M., Methven, J., Masato, G. and Chagnon, J.M. (2014) Systematic model forecast error in Rossby wave structure. *Geophysical Research Letters*, 41(8), 2979–2987.

Haiden, T., Janousek, M., Bidlot, J.-R., Ferranti, L., Prates, F., Vitart, F., Bauer, P. and Richardson, D. (2016). Evaluation of ECMWF forecasts, including the 2016 resolution upgrade. Technical Memorandum no. 792, Reading, UK: ECMWF.

Hardiman, S.C., Boutle, I.A., Bushell, A.C., Butchart, N., Cullen, M.J.P., Field, P.R., Furtado, K., Manners, J.C., Milton, S.F., Morcrette, C.J., O'Connor, F.M., Shipway, B.J., Smith, C., Walters, D.N., Willett, M.R., Williams, K.D., Wood, N., Abraham, N.L., Keeble, J., Maycock, A.C., Thuburn, J. and Woodhouse, M.T.

(2015) Processes controlling tropical tropopause temperature and stratospheric water vapor in climate models. *Journal of Climate*, 28(16), 6516–6535.

Harvey, B.J., Methven, J. and Ambaum, M.H.P. (2016) Rossby wave propagation on potential vorticity fronts with finite width. *Journal of Fluid Mechanics*, 794, 775–797.

Harvey, B.J., Methven, J. and Ambaum, M.H.P. (2018) An adiabatic mechanism for the reduction of jet meander amplitude by potential vorticity filamentation. *Journal of the Atmospheric Sciences*, 75(12), 4091–4106.

Harvey, B.J., Methven, J., Sanchez, C. and Schäfler, A. (2020) Diabatic generation of negative potential vorticity and its impact on the North Atlantic jet stream. *Quarterly Journal of the Royal Meteorological Society*, 146(728), 1477–1497.

Hegglin, M.I., Boone, C.D., Manney, G.L. and Walker, K.A. (2009) A global view of the extratropical tropopause transition layer from Atmospheric Chemistry Experiment Fourier Transform Spectrometer O_3, H_2O, and CO. *Journal of Geophysical Research: Atmospheres*, 114(D7). https://doi.org/10.1029/2008JD009984.

Hoskins, B.J. and James, I.N. (2014) *Fluid Dynamics of the Midlatitude Atmosphere*. Oxford, UK: Wiley Blackwell.

Ingleby, N.B. (2017). An assessment of different radiosonde types 2015/2016. Technical Memorandum No. 807, Reading, UK: ECMWF.

Ingleby, N.B. and Edwards, D. (2015) Changes to radiosonde reports and their processing for numerical weather prediction. *Atmospheric Science Letters*, 16(1), 44–49.

Ingleby, N.B., Lorenc, A.C., Ngan, K., Rawlins, R. and Jackson, D. (2012). Improved variational analyses using a nonlinear humidity control variable: formulation and trials. Foundation and Weather Science Technical Report no. 558, Exeter, UK: Met Office.

Ingleby, N.B., Pauley, P., Kats, A., Ator, J., Keyser, D., Doerenbecher, A., Fucile, E., Hasegawa, J., Toyoda, E., Kleinert, T., Qu, W., St. James, J., Tennant, W. and Weedon, R. (2016) Progress toward high-resolution, real-time radiosonde reports. *Bulletin of the American Meteorological Society*, 97(11), 2149–2161.

Jauhiainen, H., Survo, P., Lehtinen, R., Lentonen, J. and Oyj, V. (2014). Radiosonde RS41 and RS92 key differences and comparison test results in different locations and climates, in Proceedings of TECO-2014, WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observations, 7–9 July, St. Petersberg, Russian Federation. Geneva, Switzerland: WMO.

Jensen, M.P., Holdridge, D., Survo, P., Lehtinen, R., Baxter, S., Toto, T. and Johnson, K.L. (2016) Comparison of Vaisala radiosondes RS41 and RS92 at the ARM Southern Great Plains site. *Atmospheric Measurement Techniques*, 9(7), 3115–3129.

Joshi, M.M., Charlton, A.J. and Scaife, A.A. (2006) On the influence of stratospheric water vapor changes on the tropospheric circulation. *Geophysical Research Letters*, 33(9). https://doi.org/10.1029/2006GL025983.

Kunz, A., Spelten, N., Konopka, P., Müller, R., Forbes, R.M. and Wernli, H. (2014) Comparison of Fast In situ Stratospheric Hygrometer (FISH) measurements of water vapor in the upper troposphere and lower stratosphere (UTLS) with ECMWF (re)analysis data. *Atmospheric Chemistry and Physics*, 14, 10803–10822.

Malardel, S., Wedi, N., Deconinck, W., Diamantakis, M., Kühnlein, C., Mozdzynski, G., Hamrud, M. and Smolarkiewicz, P. (2016) A new grid for the IFS. *ECMWF Newsletter*, 146, 23–28.

Maycock, A.C., Shine, K.P. and Joshi, M.M. (2011) The temperature response to stratospheric water vapour changes. *Quarterly Journal of the Royal Meteorological Society*, 137, 1070–1082.

Maycock, A.C., Joshi, M.M., Shine, K.P. and Scaife, A.A. (2013) The circulation response to idealized changes in stratospheric water vapor. *Journal of Climate*, 26(2), 545–561.

Motl, M. (2014) Vaisala RS41 trial in the Czech Republic. *Vaisala News*, 192/2014, 14–17.

Nash, J., Oakley, T., Vömel, H. and Wei, L. (2010). WMO intercomparison of high-quality radiosonde systems. Instruments and Observing Methods, Report no. 107, Geneva, Switzerland: World Meteorological Organisation.

Oh, J., Son, S.-W., Williams, K., Walters, D., Kim, J., Willett, M., Earnshaw, P., Bushell, A., Kim, Y. and Kim, J. (2018) Ozone sensitivity of tropical upper-troposphere and stratosphere temperature in the MetOffice Unified Model. *Quarterly Journal of the Royal Meteorological Society*, 144, 2001–2009.

Oikonomou, E.K. and O'Neill, A. (2006) Evaluation of ozone and water vapor fields from the ECMWF reanalysis ERA-40 during 1991–99 in comparison with UARS satellite and MOZAIC aircraft observations. *Journal of Geophysical Research: Atmospheres*, 111(D14). https://doi.org/10.1029/2004JD005341.

Pope, V.D., Pamment, J.A., Jackson, D.R. and Slingo, A. (2001) The representation of water vapor and its dependence on vertical resolution in the Hadley Centre climate model. *Journal of Climate*, 14(14), 3065–3085.

Randel, W.J., Wu, F. and Forster, P. (2007) The extratropical tropopause inversion layer: global observations with GPS data, and a radiative forcing mechanism. *Journal of the Atmospheric Sciences*, 64(12), 4489–4496.

Saffin, L., Gray, S.L., Methven, J. and Williams, K.D. (2017) Processes maintaining tropopause sharpness in numerical models. *Journal of Geophysical Research: Atmospheres*, 122(18), 9611–9627.

Schäfler, A., Craig, G., Wernli, H., Arbogast, P., Doyle, J.D., McTaggart-Cowan, R., Methven, J., Rivière, G., Ament, F., Boettcher, M., Bramberger, M., Cazenave, Q., Cotton, R., Crewell, S., Delanoë, J., Dörnbrack, A., Ehrlich, A., Ewald, F., Fix, A., Grams, C.M., Gray, S.L., Grob, H., Gross, S., Hagen, M., Harvey, B., Hirsch, L., Jacob, M., Kölling, T., Konow, H., Lemmerz, C., Lux, O., Magnusson, L., Mayer, B., Mech, M., Moore, R., Pelon, J., Quinting, J., Rahm, S., Rapp, M., Rautenhaus, M., Reitebuch, O., Reynolds, C.A., Sodemann, H., Spengler, T., Vaughan, G., Wendisch, M., Wirth, M., Witschas, B., Wolf, K. and Zinner, T. (2018) The North Atlantic Waveguide and Downstream Impact EXperiment. *Bulletin of the American Meteorological Society*, 99(8), 1607–1637.

Schäfler, A., Harvey, B., Methven, J., Doyle, J.D., Rahm, S., Reitebuch, O., Weiler, F. and Witschas, B. (2020) Observation of jet stream winds during NAWDEX and characterization of systematic meteorological analysis errors. *Monthly Weather Review*, 148(7), 2889–2907.

Shepherd, T.G., Polichtchouk, I., Hogan, R.J. and Simmons, A.J. (2018). Report on Stratosphere Task Force. Technical Memorandum 824, Reading, UK: ECMWF.

Shine, K.P. and Myhre, G. (2020) The spectral nature of stratospheric temperature adjustment and its application to halocarbon radiative forcing. *Journal of Advances in Modeling Earth Systems*, 12(3). https://doi.org/10.1029/2019MS001951.

Smith, C.A., Haigh, J.D. and Toumi, R. (2001) Radiative forcing due to trends in stratospheric water vapour. *Geophysical Research Letters*, 28(1), 179–182.

Solomon, S., Rosenlof, K.H., Portmann, R.W., Daniel, J.S., Davis, S.M., Sanford, T.J. and Plattner, G.-K. (2010) Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science*, 327(5970), 1219–1223.

Sonntag, D. (1990) Important new values of the physical constants of 1986, vapour pressure formulations based on the ITS-90, and psychrometer formulae. *Zeitschrift für Meteorologie*, 70(5), 340–344.

Sonntag, D. (1994) Advancements in the field of hygrometry. *Meteorologische Zeitschrift*, 3, 51–66.

Stenke, A., Grewe, V. and Ponater, M. (2008) Lagrangian transport of water vapor and cloud water in the ECHAM4 GCM and its impact on the cold bias. *Climate Dynamics*, 31(5), 491–506.

Tandon, N.F., Polvani, L.M. and Davis, S.M. (2011) The response of the tropospheric circulation to water vapor-like forcings in the stratosphere. *Journal of Climate*, 24(21), 5713–5720.

Teixeira, J. and AIRS Science Team (2013) *AIRS/Aqua L3 Daily Support Product (AIRS-only) 1 degree x 1 degree V006*. Greenbelt, MD: Goddard Earth Sciences Data and Information Services Center (GES DISC).

Vaisala (2013). *Vaisala Radiosonde RS92-SGP*. Vantaa, Finland. Technical Datasheet B210358EN-F. Available at: https://www.vaisala.com/sites/default/files/documents/RS92SGP-Datasheet-B210358EN-F-LOW.pdf; accessed 17 August 2021.

Vaisala (2014). *Comparison of Vaisala Radiosondes RS41 and RS92*. Vantaa, Finland. White Paper Datasheet B211317EN-B. Available at: https://www.vaisala.com/sites/default/files/documents/RS-Comparison-White-Paper-B211317EN.pdf; accessed 17 August 2021.

Vaisala (2017a). Vaisala Radiosonde RS41 Measurement Performance. Vantaa, Finland. White Paper Datasheet B211356EN-B. Available at: https://www.vaisala.com/sites/default/files/documents/WEA-MET-RS41-Performance-White-paper-B211356EN-B-LOW-v3.pdf. Accessed 17 August 2021.

Vaisala (2017b). Vaisala Radiosonde RS41?SGP. Vantaa, Finland. Technical Datasheet B211444EN?E. Available at: https://www.vaisala.com/sites/default/files/documents/WEA-MET-RS41SGP-Datasheet-B211444EN.pdf. Accessed 17 August 2021.

Vaisala (2020). *Vaisala RS92 Radiosonde Technical Changes*. Vantaa, Finland. Available at: https://www.vaisala.com/sites/default/files/documents/Vaisala%20Radiosonde%20RS92%20technical%20changes_0.pdf; accessed 23 October 2020.

Van Thien, L., Gallus, W.A., Olsen, M.A. and Livesey, N. (2010) Comparison of aura MLS water vapor measurements with GFS and NAM analyses in the upper troposphere–lower stratosphere. *Journal of Atmospheric and Oceanic Technology*, 27(2), 274–289.

Walters, D., Boutle, I.A., Brooks, M., Melvin, T., Stratton, R.A., Vosper, S.B., Wells, H., Williams, K., Wood, N., Allen, T., Bushell, A., Copsey, D., Earnshaw, P., Edwards, J., Gross, M., Hardiman, S., Harris, C., Heming, J., Klingaman, N., Levine, R., Manners, J., Martin, G., Milton, S.B., Mittermaier, M., Morcrette, C.J., Riddick, T., Roberts, M., Sanchez, C., Selwood, P., Stirling, A., Smith, C., Suri, D., Tennant, W., Vidale, P.L., Wilkinson, J., Willett, M., Woolnough, S. and Xavier, P. (2017) The Met Office Unified Model global atmosphere 6.0/6.1 and JULES global land 6.0/6.1 configurations. *Geoscientific Model Development*, 10(4), 1487–1520.

Wang, J., Zhang, L., Dai, A., Immler, F., Sommer, M. and Vömel, H. (2013) Radiation dry bias correction of Vaisala RS92 humidity data and its impacts on historical radiosonde data. *Journal of Atmospheric and Oceanic Technology*, 30(2), 197–214.

WMO (1957) Meteorology – a three-dimensional science. *WMO Bulletin*, 6(4), 134–139.

Woiwode, W., Dörnbrack, A., Polichtchouk, I., Johansson, S., Harvey, B., Höpfner, M., Ungermann, J. and Friedl-Vallon, F. (2020) Technical note: Lowermost-stratospheric moist bias in ECMWF IFS model diagnosed from airborne GLORIA observations during winter–spring 2016. *Atmospheric Chemistry and Physics*, 20, 15379–15387.

World Ozone and Ultraviolet Radiation Data Centre (2020). WOUDC OzoneSonde Dataset. Available at: https://woudc.org/data/dataset_info.php?id=ozonesonde; accessed 17 August 2021.