



# Evaluating the performance of a Picarro G2207-*i* analyser for high-precision atmospheric O<sub>2</sub> measurements

Leigh S. Fleming<sup>1,a</sup>, Andrew C. Manning<sup>1</sup>, Penelope A. Pickers<sup>1</sup>, Grant L. Forster<sup>1,2</sup>, and Alex J. Etchells<sup>1</sup>

<sup>1</sup>Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK

<sup>2</sup>National Centre for Atmospheric Science, University of East Anglia, Norwich, UK

<sup>a</sup>now at: GNS Science, Gracefield, Lower Hutt, 5040, New Zealand

**Correspondence:** Leigh S. Fleming (l.fleming@gns.cri.nz)

Received: 8 July 2022 – Discussion started: 19 July 2022

Revised: 17 October 2022 – Accepted: 29 October 2022 – Published: 24 January 2023

**Abstract.** Fluxes of oxygen (O<sub>2</sub>) and carbon dioxide (CO<sub>2</sub>) in and out of the atmosphere are strongly coupled for terrestrial biospheric exchange processes and fossil fuel combustion but are uncoupled for oceanic air–sea gas exchange. High-precision measurements of both species can therefore provide constraints on the carbon cycle and can be used to quantify fossil fuel CO<sub>2</sub> (ffCO<sub>2</sub>) emission estimates. In the case of O<sub>2</sub>, however, due to its large atmospheric mole fraction (~20.9%) it is very challenging to measure small variations to the degree of precision and accuracy required for these applications. We have tested an atmospheric O<sub>2</sub> analyser based on the principle of cavity ring-down spectroscopy (Picarro Inc., model G2207-*i*), both in the laboratory and at the Weybourne Atmospheric Observatory (WAO) field station in the UK, in comparison to well-established, pre-existing atmospheric O<sub>2</sub> and CO<sub>2</sub> measurement systems.

In laboratory tests analysing dry air in high-pressure cylinders, we found that the best precision was achieved with 30 min averaging and was ±0.5 ppm (~±2.4 per meg). Also from continuous measurements from a cylinder of dry air, we found the 24 h peak-to-peak range of hourly averaged values to be 1.2 ppm (~5.8 per meg). These results are close to atmospheric O<sub>2</sub> compatibility goals as set by the UN World Meteorological Organization. However, from measurements of ambient air conducted at WAO we found that the built-in water correction of the G2207-*i* does not sufficiently correct for the influence of water vapour on the O<sub>2</sub> mole fraction. When sample air was dried and a 5-hourly baseline correction with a reference gas cylinder was employed, the G2207-*i*'s results showed an average difference from the established O<sub>2</sub> analyser of 13.6 ± 7.5 per meg (over 2 weeks of conti-

nuous measurements). Over the same period, based on measurements of a so-called “target tank”, analysed for 12 min every 7 h, we calculated a repeatability of ±5.7 ± 5.6 per meg and a compatibility of ±10.0 ± 6.7 per meg for the G2207-*i*. To further examine the G2207-*i*'s performance in real-world applications we used ambient air measurements of O<sub>2</sub> together with concurrent CO<sub>2</sub> measurements to calculate ffCO<sub>2</sub>. Due to the imprecision of the G2207-*i*, the ffCO<sub>2</sub> calculated showed large differences from that calculated from the established measurement system and had a large uncertainty of ±13.0 ppm, which was roughly double that from the established system (±5.8 ppm).

## 1 Introduction

Oxygen (O<sub>2</sub>) is the most abundant molecule in the atmosphere after nitrogen (N<sub>2</sub>), with an atmospheric background mole fraction of approximately 20.9%. Due to this large atmospheric background, O<sub>2</sub> measurements are sensitive to variations in the mole fractions of other atmospheric species, such as carbon dioxide (CO<sub>2</sub>), due to dilution effects. O<sub>2</sub> measurements are therefore typically reported on a relative scale calculated as the change in the ratio of O<sub>2</sub> to N<sub>2</sub> relative to a standard O<sub>2</sub>/N<sub>2</sub> ratio, as given in Eq. (1), and expressed in “per meg” units.

$$\delta \left( \frac{\text{O}_2}{\text{N}_2} \right) = \left( \frac{\text{O}_2/\text{N}_2_{\text{sample}} - \text{O}_2/\text{N}_2_{\text{reference}}}{\text{O}_2/\text{N}_2_{\text{reference}}} \right) \times 10^6 \quad (1)$$

In practice, atmospheric N<sub>2</sub> is far less variable than O<sub>2</sub>, meaning that changes in the O<sub>2</sub>/N<sub>2</sub> ratio can be assumed to

be representative of O<sub>2</sub> mole fraction (Keeling and Shertz, 1992). In comparing changes in O<sub>2</sub> to changes in CO<sub>2</sub>, on a mole for mole basis, a 1 per meg change in O<sub>2</sub> is equivalent to a 0.2094 ppm (parts per million) change in CO<sub>2</sub> mole fraction (Keeling et al., 1998).

Over the past 3 decades, atmospheric O<sub>2</sub> has been decreasing at an average rate of  $\sim 15$  per meg per year, primarily owing to fossil fuel combustion (Keeling and Manning, 2014); over the same period, atmospheric CO<sub>2</sub> has been increasing at an average rate of 2 ppm yr<sup>-1</sup> (Dlugokencky and Tans, 2022), also predominantly due to fossil fuel combustion. For most processes that cause variability in atmospheric O<sub>2</sub>, there is an anti-correlated change in atmospheric CO<sub>2</sub>; therefore, high-precision measurements of atmospheric O<sub>2</sub> play an increasingly important role in our understanding of atmospheric CO<sub>2</sub>, carbon cycling, and other biogeochemical processes (e.g. Pickers et al., 2017; Resplandy et al., 2019; Battle et al., 2019; Tohjima et al., 2019). Fluxes of O<sub>2</sub> and CO<sub>2</sub> in and out of the atmosphere are strongly coupled for terrestrial biosphere exchange with a global average oxidative ratio (OR) in the range of 1.03 to 1.10 mol mol<sup>-1</sup> (Severinghaus, 1995). For fossil fuel combustion, dependent on fuel type, the OR is in the range of 1.17 to 1.95 mol mol<sup>-1</sup> (Keeling, 1988b). Whereas O<sub>2</sub> and CO<sub>2</sub> fluxes are uncoupled for oceanic air–sea gas exchange primarily due to inorganic reactions in the water involving the carbonate system and not O<sub>2</sub>, as well as differences in air–sea equilibration times between the two gases.

The relationship between O<sub>2</sub> and CO<sub>2</sub> fluxes has also allowed for the derivation of the tracer “atmospheric potential oxygen” (APO), as defined in Eq. (2) (Stephens et al., 1998).

$$\text{APO} = \text{O}_2 + (-1.1 \times \text{CO}_2), \quad (2)$$

where the factor  $-1.1$  represents the mean value of the O<sub>2</sub> : CO<sub>2</sub> OR for terrestrial biosphere photosynthesis and respiration (Severinghaus, 1995), and where we have ignored very minor influences from methane and carbon monoxide. APO is therefore, by definition, invariant with respect to the terrestrial biosphere. Changes in APO therefore mainly reflect changes in ocean–atmosphere exchange of O<sub>2</sub> and CO<sub>2</sub> (primarily on seasonal and longer timescales), with a contribution from fossil fuels on both shorter and longer timescales. APO can thus be used to examine oceanic CO<sub>2</sub> fluxes and to quantify fossil fuel CO<sub>2</sub> (ffCO<sub>2</sub>) emissions (Pickers et al., 2022).

The World Meteorological Organization (WMO) Global Atmosphere Watch (GAW) programme has established a measurement compatibility goal for O<sub>2</sub> of  $\pm 2$  per meg ( $\pm 0.4$  ppm) (Crotwell et al., 2019), where compatibility refers to the acceptable level of agreement between two field stations or laboratories when measuring the same air sample. This is the scientifically desirable level of compatibility required to resolve, for example, latitudinal gradients and long-term trends (Crotwell et al., 2019). There is also a WMO extended compatibility goal of  $\pm 10$  per meg ( $\pm 2$  ppm), which

is suitable for some specific applications when expected variations are relatively large (Crotwell et al., 2019), such as fossil fuel quantification in large cities. In order to meet the WMO compatibility goals, it is recommended that a measurement system’s repeatability should not exceed half of the compatibility goal (i.e.  $\pm 1$  per meg;  $\pm 0.2$  ppm). Repeatability refers to the closeness of agreement between results of repeated measurements of the same measure (which is also sometimes referred to as the measurement system’s precision). However, routinely achieving a measurement repeatability of  $\pm 1$  per meg is not achievable for almost any laboratories or field stations making high-precision measurements of atmospheric O<sub>2</sub>. The large atmospheric background of O<sub>2</sub> makes it extremely challenging to measure the relatively small variations to the level of repeatability required, since measuring a change of, for example, 0.2 ppm against the background ( $\sim 209\,400$  ppm) requires a relative repeatability of 0.0001 %.

Presently, there are several different analytical techniques available for measuring atmospheric O<sub>2</sub> to high precision: interferometry (Keeling, 1988a), isotope ratio mass spectrometry (Bender et al., 1994), paramagnetic techniques (Manning et al., 1999), vacuum ultraviolet absorption (VUV; Stephens et al., 2011), gas chromatography (Tohjima, 2000), and electrochemical fuel cells (Stephens et al., 2007). The most precise of these current methods is the VUV absorption technique; however, VUV O<sub>2</sub> analysers are “homemade” and are not commercially available, thus limiting their widespread application. None of these techniques are “off-the-shelf” systems, all of them are complex and time-consuming systems to design, build, and optimise, with very precise pressure, temperature, and flow control needed. All of the techniques also require frequent interruption to sample measurement to carry out calibration procedures (Kozlova and Manning, 2009). The supply of calibration gases for such systems is particularly labour intensive, due to both their relatively rapid consumption rate and the fact that no commercial gas supply company is able to provide suitable gas mixtures for atmospheric O<sub>2</sub> research. Accurate, high-precision atmospheric O<sub>2</sub> measurements therefore remain challenging. An alternative commercially available O<sub>2</sub> analyser with fewer requirements for external gas handling, air-sample drying, and calibration procedures could consequently advance the field of atmospheric O<sub>2</sub> measurements if the required performance could be achieved and if it were relatively easy to operate with low maintenance requirements and a lower rate of calibration gas consumption.

In this paper we present the results from the analysis of a Picarro Inc. G2207-*i* oxygen analyser, which operates on the principle of cavity ring-down spectroscopy technology (CRDS) (hereafter referred to as the G2207-*i*) and evaluate its performance in comparison to established O<sub>2</sub> measurement systems in the University of East Anglia (UEA) Carbon Related Atmospheric Measurements (CRAM) Laboratory and at the Weybourne Atmospheric Observatory (WAO);

North Norfolk, UK). Unlike most other analytical techniques used for atmospheric O<sub>2</sub> measurements, it is intended that the G2207-*i* should not require a continuous reference gas supply, and it has built-in pressure and flow control and the potential for greatly reduced sample drying requirements due to a built-in water measurement and correction procedure. These features make the G2207-*i* a potentially desirable analyser for high-precision atmospheric O<sub>2</sub> research, but we note that it would still require the same rigorous calibration procedures as other analysers (Kozlova and Manning, 2009), albeit possibly at reduced frequency. In this paper we quantify the compatibility, repeatability, and drift rates in the context of the WMO/GAW guidelines (Crotwell et al., 2019). In order to further examine the performance of the G2207-*i* in real-world applications, we also calculated ffCO<sub>2</sub> from concurrent O<sub>2</sub> and CO<sub>2</sub> measurements, using the novel methodology presented by Pickers et al. (2022). We compare ffCO<sub>2</sub> calculated with O<sub>2</sub> measurements from the G2207-*i* installed at WAO with ffCO<sub>2</sub> calculated from the established O<sub>2</sub> system employing a Sable Systems International Inc. “Oxzilla II” fuel cell analyser.

## 2 Methods

### 2.1 Picarro G2207-*i* O<sub>2</sub> analyser

The Picarro G2207-*i* O<sub>2</sub> analyser measures the mole fractions of the two most abundant atmospheric O<sub>2</sub> isotopologues, <sup>16</sup>O<sup>16</sup>O and <sup>16</sup>O<sup>18</sup>O, through absorption spectra at 7882.18670 and 7882.050155 cm<sup>-1</sup>, respectively (Berhanu et al., 2019). The design principles of this analyser have been described in detail by Berhanu et al. (2019). In our study we evaluate only what is called the “O<sub>2</sub> concentration” mode, measuring only the <sup>16</sup>O<sup>16</sup>O isotopologue. In the other mode, called the “δ<sup>18</sup>O plus O<sub>2</sub> concentration” mode, O<sub>2</sub> mole fraction values are considerably less precise, as the analyser is not optimised for <sup>16</sup>O<sup>16</sup>O measurements (primarily via a different set point for the pressure in the cavity). The analyser reports both “wet” and “dry” O<sub>2</sub> mole fraction values. The wet values (O<sub>2,NC</sub>; NC stands for “not corrected”) do not have any correction applied to them, whereas the dry values (O<sub>2,WC</sub>; WC stands for “water corrected”) are corrected for the dilution effect of water vapour on the O<sub>2</sub> mole fraction, as well as spectroscopic interference, using the analyser’s parallel water vapour mole fraction measurements. The G2207-*i* data sheet states a measurement precision of 5 ppm +0.1 % of the reading (1-σ, 5 s) for the water vapour mole fraction.

### 2.2 CRAM laboratory measurement of cylinder gases

The performance of the G2207-*i* was evaluated in the UEA CRAM Laboratory by measuring a suite of 12 gas cylinders all containing dry natural air with varying O<sub>2</sub> mole fractions. The cylinders were stored horizontally in a thermally insu-

lated “Blue Box” enclosure in order to prevent gravitational and thermal fractionation of O<sub>2</sub> relative to N<sub>2</sub> (Keeling et al., 2007). The O<sub>2</sub> composition of each of these cylinders was precisely defined on the Scripps Institution of Oceanography (SIO) O<sub>2</sub> scale (Keeling et al., 2007) using a VUV O<sub>2</sub> analyser, which is also in the CRAM Laboratory. The CO<sub>2</sub> mole fraction was defined on the “WMO CO<sub>2</sub> X2007” scale (Zhao and Tans, 2006) using a Siemens Corp. Ultramat model 6F non-dispersive infrared (NDIR) CO<sub>2</sub> analyser. Five of these cylinders were working secondary standards (WSSes), which were used to calibrate the G2207-*i*, one was a reference tank (RT; explained below in Sect. 2.3.2), while the other six were treated as cylinders with unknown mole fractions (Table 1). The six “unknown” cylinders were used to evaluate the performance of the analyser with a CO<sub>2</sub> mole fraction range of 375 to 443 ppm and an O<sub>2</sub>/N<sub>2</sub> ratio range of -915 to +435 per meg, a much larger range than would typically be observed in ambient air.

The cylinders were run consecutively, starting with the six “unknowns” and ending with the five WSSes, with the RT run at the beginning and end; this sequence was repeated twice. Each of the gas cylinders was flushed for 20 min prior to running on the G2207-*i* to allow for removal of stagnant air and equilibration of the pressure regulators; air from each cylinder was then passed through the analyser for 20 min, with the first 8 min of data discarded to allow flushing of the previous cylinder’s air from the cavity and to maintain consistency with the flushing time employed in subsequent WAO tests (Sect. 2.3.2). The remaining 12 min for each cylinder was then averaged to give the “raw” O<sub>2,NC</sub> value for each cylinder as measured on the G2207-*i*.

The G2207-*i* has a linear response to O<sub>2</sub> mole fraction (Eq. 3)

$$y = Bx + C \quad (3)$$

where *B* and *C* are the coefficients derived from the slope and intercept of the linear regression calculated from the measurement of the WSSes. Therefore, a minimum of two WSS cylinders are required to determine the *B* and *C* coefficients, but by using five we are able to calculate the coefficient of determination (*R*<sup>2</sup>), as well as providing more robustness in the fit. The calibration equation was used to convert the “raw” O<sub>2,NC</sub> values taken from the G2207-*i* (*x* in Eq. 3) into what we call “ppm equivalent” (ppmEquiv) O<sub>2</sub> units (*y* in Eq. 3), as described in Kozlova and Manning (2009). A linear interpolation between the RT at the beginning and end of each run was used as a baseline for the run and subtracted from all other cylinder measurements to correct for short-term analyser variations. The calibration curve (Eq. 3) for the G2207-*i* was also determined relative to the interpolated RT values (WSS – RT); thus, all the unknown cylinder measurements could be converted into ppmEquiv. The ppmEquiv O<sub>2</sub> units were then converted to per meg units, providing a δ(O<sub>2</sub>/N<sub>2</sub>)

**Table 1.** Declared O<sub>2</sub>/N<sub>2</sub> ratios and CO<sub>2</sub> mole fractions with ±1σ standard deviations of the five WSSes, RT, and six “unknown” cylinder gases used in the CRAM Laboratory tests of the G2207-*i*.

Cylinder number	Cylinder ID	Declared O <sub>2</sub> (per meg) <sup>a</sup>	Declared CO <sub>2</sub> (ppm) <sup>b</sup>
WSS1	D089507	−565.5 ± 1.3	428.741 ± 0.018
WSS2	D801299	−486.1 ± 3.0	381.230 ± 0.016
WSS3	D073409	−658.4 ± 2.2	398.875 ± 0.018
WSS4	D073419	−926.4 ± 5.9	440.355 ± 0.072
WSS5	D073418	−782.7 ± 5.6	413.662 ± 0.057
RT	CC78691	−414.3 ± 0.8	384.915 ± 0.005
1	D273555	−914.8 ± 0.7	443.384 ± 0.013
2	D399093	−880.5 ± 0.9	415.246 ± 0.003
3	ND29112	−582.0 ± 1.0	399.976 ± 0.004
4	ND29110	−375.0 ± 1.3	381.544 ± 0.004
5	D273559 <sup>c</sup>	411.7 ± 2.1	375.122 ± 0.007
6	D801298 <sup>c</sup>	434.6 ± 0.3	412.934 ± 0.002

<sup>a</sup> Values declared with a VUV O<sub>2</sub> analyser in the CRAM laboratory traceable to the SIO O<sub>2</sub> scale.

<sup>b</sup> Values declared with a Siemens Ultramat 6F NDIR CO<sub>2</sub> analyser in the CRAM Laboratory traceable to the WMO CO<sub>2</sub> X2007 scale.

<sup>c</sup> The O<sub>2</sub> values of these cylinders are far outside the range observed in ambient air, thus are less relevant to the applications of atmospheric observations but have been included in this analysis for completeness of examining the analyser’s performance.

value for each unknown cylinder, using Eq. (4).

$$\delta \left( \frac{\text{O}_2}{\text{N}_2} \right) = \frac{\delta \text{O}_2 + (\text{CO}_2 - 363.29) \times S_{\text{O}_2}}{S_{\text{O}_2} \times (1 - S_{\text{O}_2})} \quad (4)$$

where, δO<sub>2</sub> is the calibrated G2207-*i* O<sub>2,NC</sub> value in ppmE-equiv units, CO<sub>2</sub> is the declared cylinder CO<sub>2</sub> mole fraction from the Siemens analyser in ppm, S<sub>O<sub>2</sub></sub> is 0.2094, which is the standard mole fraction of O<sub>2</sub> molecules in dry air (Tohjima et al., 2005), and 363.29 is an arbitrary CO<sub>2</sub> reference value in ppm, inherent to the SIO O<sub>2</sub> scale (Stephens et al., 2007).

### 2.3 Weybourne Atmospheric Observatory field tests

Weybourne Atmospheric Observatory (WAO) is located on the north Norfolk coast, UK (52°57′02″ N, 1°07′19″ E), approximately 35 km north-northwest of Norwich, 170 km northeast of London, and 200 km east of Birmingham. It is part of the European Union’s Integrated Carbon Observation System (ICOS) and the World Meteorological Organization’s (WMO) Global Atmosphere Watch (GAW) programme. High-precision, high-accuracy, continuous measurements of a wide array of atmospheric gas species (including greenhouse gases, isotopes, reactive gases) are carried out at a fine temporal scale, funded in part through the UK’s National Centre for Atmospheric Science (NCAS) long-term measurement programme.

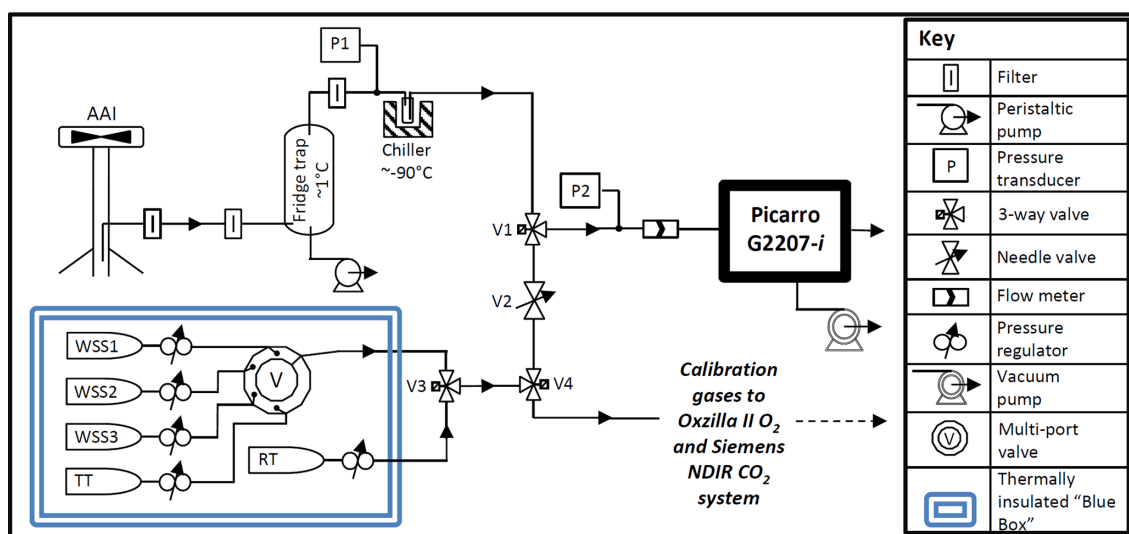
Atmospheric O<sub>2</sub> and CO<sub>2</sub> have been measured continuously at WAO since 2008 (Wilson, 2013). O<sub>2</sub> is measured with an Oxzilla II O<sub>2</sub> analyser (Sable Systems International Inc.) (hereafter referred to as the “Oxzilla”), and CO<sub>2</sub> is measured with an Ultramat 6E NDIR analyser (Siemens Corp.). These analysers are arranged in series, with the air sample

first passing through the Ultramat 6E and then the Oxzilla, with rigorous gas handling and calibration protocols followed (as in Stephens et al., 2007).

The G2207-*i* was installed at WAO from 23 October 2019–2 November 2019, sampling from a solar shield aspirated air inlet (AAI) at a height of 10 m above ground level (a.g.l.; 20 m above sea level, a.s.l.). The AAI protects the inlet from solar radiation and generates a continuous air flow over the inlet, thus preventing the differential fractionation of O<sub>2</sub> molecules relative to N<sub>2</sub> molecules due to ambient temperature variations (Blaine et al., 2006) and relatively slow inlet flow rates (Manning, 2001). A diagram of the gas handling set-up for the G2207-*i* at WAO is displayed in Fig. 1.

#### 2.3.1 Drying

Water vapour mole fractions in the troposphere vary from a few parts per million to a few percent over small temporal and spatial scales. This water vapour has a diluting effect on atmospheric gas measurement. A 1 ppm increase of water vapour will dilute the measured atmospheric O<sub>2</sub> by approximately 1.3 per meg (Stephens et al., 2007); thus, the existing method for high-precision atmospheric O<sub>2</sub> measurements is to dry the sample air to less than 1 ppm water vapour content before measurement. All calibration and RT gases are also dried to less than 1 ppm water vapour. Furthermore, measurements using spectroscopic techniques are also sensitive to water vapour variability due to changes in the degree of pressure broadening of the spectroscopic lines used to measure the O<sub>2</sub> and δ<sup>18</sup>O<sub>2</sub>. Water vapour correction has previously been successfully implemented for measurements of CO<sub>2</sub> and methane (CH<sub>4</sub>) with CRDS analysers (Chen et al., 2010); however, in order to achieve accuracies within the



**Figure 1.** Gas handling diagram of the Picarro G2207-*i* installed at WAO. (AAI, aspirated air inlet; WSS, working secondary standard; RT, reference tank; TT, target tank). Calibration gases were shared with the established O<sub>2</sub> and CO<sub>2</sub> system (using V4), but the established system has its own AAI, pump, drying system, and pressure and flow control (not depicted here).

WMO goal of 1 % H<sub>2</sub>O custom coefficients must be obtained for each analyser (Rella et al., 2013).

As discussed in Sect. 2.1, O<sub>2</sub> measurements are reported by the G2207-*i* as “wet” (O<sub>2,NC</sub>) and, after the implementation of water correction, “dry” (O<sub>2,WC</sub>). In order to evaluate the effectiveness of the built-in water correction procedure for compensating for water vapour dilution, ambient air was sampled with three different drying regimes: no drying, partial drying, and full drying. Under the full-drying conditions (which is the current standard practice), the sample air passed through a fridge trap ( $\sim 1^\circ\text{C}$ ) and a cryogenic chiller trap ( $\sim -90^\circ\text{C}$ ), removing water vapour to  $< 1$  ppm. Under partial drying the chiller was bypassed, so the sample air only passed through the fridge trap, which dries the air to approximately 5000 ppm of water vapour. With no drying, both the chiller and fridge were bypassed. Air was simultaneously sampled through a separate AAI (10 m a.g.l.) into the pre-existing O<sub>2</sub> and CO<sub>2</sub> system with full drying during each of these stages. The time difference between air travelling from the AAIs to each of the two analysers was accounted for.

To evaluate the built-in water correction procedure of the G2207-*i*, the O<sub>2,WC</sub> values were compared with measurements from the Oxzilla (which was continuously sampling fully dried air) for the no drying and partial drying periods, and the O<sub>2,NC</sub> and O<sub>2,WC</sub> G2207-*i* values were compared to the Oxzilla when sampling fully dried air.

### 2.3.2 Calibration procedure

A tailor-made calibration protocol was developed for the G2207-*i* following ICOS atmospheric station specifications (ICOS-RI, 2020). The calibration cylinders were stored horizontally in a thermally insulated “Blue Box” enclosure in or-

der to prevent gravitational and thermal fractionation of O<sub>2</sub> and N<sub>2</sub>. The calibration gases consisted of three WSSes with precisely defined O<sub>2</sub> and CO<sub>2</sub> values that span the unpolluted atmospheric range (traceable to the SIO O<sub>2</sub> and WMO CO<sub>2</sub> X2007 scales) and a reference tank (RT) with O<sub>2</sub> and CO<sub>2</sub> values close to ambient air conditions at the site. The repeatability and compatibility of the analyser were evaluated using a target tank (TT) (sometimes known as a “surveillance tank”) with precisely defined O<sub>2</sub> and CO<sub>2</sub> values. With full drying of the sample air, each of the WSSes, the RT, and the TT were run for 20 minutes, of which the first 8 min was discarded due to the sweep-out time of the G2207-*i* and equilibration after valve switching and surface effects. The final 12 min were averaged to determine the cylinder value for the given run. A flushing period of 8 min and averaging time of 12 min were chosen to match that of the established system. Under partial and no drying the run time of the cylinders was increased in order to fully flush the G2207-*i* of water vapour; each cylinder was therefore run for 32 min, with the first 20 min being discarded and the final 12 min averaged.

A full three-gas WSS calibration of the G2207-*i* was run every 23 h, this frequency is intentionally not a multiple of 24 h in order to prevent aliasing the data by calibrating under environmental conditions that may occur at the same time each day. This calibration corrects for drift in the span or non-linearity of the analyser. As in the CRAM laboratory tests (see Sect. 2.2), the WSSes were used to define a calibration equation to convert the raw analyser O<sub>2</sub> values into ppmEquiv O<sub>2</sub> units. Equation (3) and the concurrent CO<sub>2</sub> measurement from the Ultramat 6E NDIR analyser were then used to convert this into per meg units.

The RT is used for data correction caused by short-term analyser drift and was run every 5 h. A linear interpolation between each of the RT run averages was treated as a baseline and subtracted from all subsequent air and cylinder measurements. The calibration curve for the G2207-*i* was also determined relative to the RT values (WSS-RT), and thus the air measurement differences can be easily converted into per meg units.

Finally, the TT was run every 7 h, this cylinder is used to quantify the repeatability and compatibility of the analyser. “Repeatability” is defined as the closeness of agreement between results of successive measurements of the same measure carried out under the same measurement conditions and is considered as a proxy for the precision of a measurement system. “Compatibility” is defined as the averaged O<sub>2</sub> value of all TT runs over time, compared to the values declared by the VUV, and provides a measure of the compatibility to the SIO scale over time (Kozlova and Manning, 2009). The TT air does not pass through the AAI or drying lines (Fig. 1), and it is therefore mainly representative of the analyser’s repeatability and compatibility only.

## 2.4 Quantifying fossil fuel CO<sub>2</sub> using atmospheric potential oxygen

In order to further assess the G2207-*i*’s performance in real-world applications, the O<sub>2,NC</sub> observations from the full-drying regime period at WAO were used to isolate the fossil fuel component of the concurrent CO<sub>2</sub> observations and then compared to the ffCO<sub>2</sub> values calculated from atmospheric potential oxygen (APO) defined from the Oxzilla O<sub>2</sub> observations following the methodology outlined in Pickers et al. (2022).

The tracer APO, derived by Stephens et al. (1998), was first calculated using Eq. (5) (using both G2207-*i* O<sub>2,NC</sub> and Oxzilla O<sub>2</sub> values); these APO values were then used to calculate ffCO<sub>2</sub> using Eq. (6).

$$\text{APO} = [\text{O}_2] + \left( \left( \frac{-1.1}{0.2094} \right) \times (350 - [\text{CO}_2]) \right), \quad (5)$$

where O<sub>2</sub> and CO<sub>2</sub> are in per meg and parts per million, respectively,  $-1.1$  is the global average O<sub>2</sub> : CO<sub>2</sub> terrestrial biosphere–atmosphere exchange ratio (Severinghaus, 1995),  $0.2094$  is the standard mole fraction of O<sub>2</sub> in dry air (Tohjima et al., 2005), and  $350$  is an arbitrary reference value for CO<sub>2</sub> (in ppm). Multiplying CO<sub>2</sub> by  $-1.1$  and dividing by  $0.2094$  converts the CO<sub>2</sub> data from parts per million to per meg units.

$$\text{ffCO}_2 = \frac{\text{APO} - \text{APO}_{\text{bg}}}{R_{\text{APO:CO}_2}}, \quad (6)$$

where APO is derived from Eq. (5) in per meg units, APO<sub>bg</sub> is the APO background, or baseline, value determined using a statistical baseline fitting procedure, and  $R_{\text{APO:CO}_2}$  is the

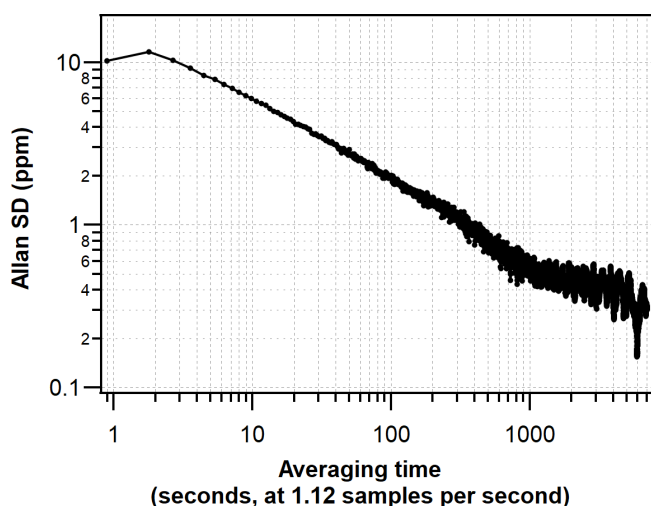
APO : CO<sub>2</sub> combustion ratio for fossil fuel emissions. The APO<sub>bg</sub> values were determined using the rfbaseline function from the IDPmisc package in R, which implements robust fitting of local regression models, with a smoothing window of 1 week (Ruckstuhl et al., 2012). The APO : CO<sub>2</sub> emission ratio ( $R_{\text{APO:CO}_2}$ ) used is  $-0.3 \text{ mol mol}^{-1}$ , an approximate mean value for WAO as determined from the COFFEE inventory (a typical value for fossil fuel emissions, given that the APO : CO<sub>2</sub> ratio = O<sub>2</sub> : CO<sub>2</sub> + 1.1; Pickers, 2016; Steinbach et al., 2011). The uncertainty on the ffCO<sub>2</sub> mole fractions was calculated using Eq. (6) with the upper and lower uncertainty limit for each variable (where the measurement uncertainty for APO was calculated by summing in quadrature the CO<sub>2</sub> and O<sub>2</sub> measurement uncertainty for each analyser) and then taking the standard deviation (SD) of the resultant ffCO<sub>2</sub> value of each combination for each hourly time stamp.

## 3 Results and discussion

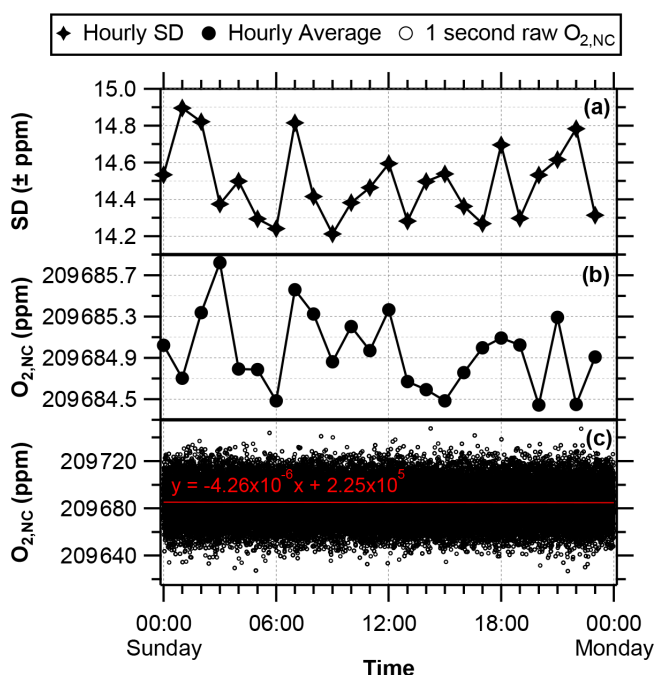
### 3.1 Precision and drift

To assess the short-term precision and optimal averaging time of the G2207-*i*, the Allan deviation technique (Werle et al., 1993) was used whilst sampling a compressed-air cylinder in the laboratory (50 L, 200 bar). The cylinder was run for 24 h with a sample flow rate of  $94 \text{ mL min}^{-1}$  and cavity pressure and temperature of 340 mbar and 45 °C, respectively. The results of this Allan deviation analysis are in agreement with those obtained by Berhanu et al. (2019), where a precision of 1 ppm ( $\sim 4.8$  per meg) was achieved after an averaging time of 300 s. Precision then continues to improve until around a 30 min averaging time where a precision of  $\sim 0.5$  ppm ( $\sim 2.4$  per meg) is reached, and it remains around that value for averaging times up to around 1 h (Fig. 2). It should be noted that unlike the hourly average and standard deviation obtained from measurement of cylinder air, the hourly averages of atmospheric data also contain natural variability in addition to analyser-related noise and drift.

To evaluate the analyser drift (i.e. the changing sensitivity of the analysers response with time), O<sub>2,NC</sub> values from the G2207-*i* were averaged to 1 h (Fig. 3b; reported in ppm where 1 ppm corresponds to a change of 4.8 per meg in the O<sub>2</sub>/N<sub>2</sub> ratio). The G2207-*i* data sheet states a maximum drift at STP (standard temperature and pressure) (over 24 h, peak-to-peak, 1 h internal average at 21 % O<sub>2</sub>) of  $< 6$  ppm. We found that over 24 h, the maximum peak-to-peak drift of the hourly averages is  $\sim 1.2$  ppm (approximately 5.8 per meg); this is better than stated by Picarro Inc. but does not meet the WMO compatibility goal of  $\pm 2$  per meg, as the internal drift of the analyser is greater than this goal. The standard deviation of each of these hourly averages is  $\sim 14.5$  ppm ( $\sim 69.6$  per meg) (Fig. 3a), this is caused by the large amount of analyser noise in the raw 1 s data points, spanning  $\sim 100$  ppm



**Figure 2.** Allan deviation plot displaying the precision of the G2207-*i* O<sub>2</sub> mole fraction measured from an ambient compressed-air cylinder.



**Figure 3.** O<sub>2,NC</sub> mole fractions from the G2207-*i* sampling dry compressed cylinder air over 24 h (reported in ppm, where 1 ppm corresponds to a change of 4.8 per meg in the O<sub>2</sub>/N<sub>2</sub> ratio). (a) Standard deviation of the hourly averaged values. (b) Hourly averaged O<sub>2,NC</sub>. (c) Raw 1 s O<sub>2,NC</sub> values, the red line depicts the linear regression line, with the equation and  $R^2$  value written above.

( $\sim 480$  per meg) (Fig. 3c). The overall drift over the 24 h of raw data however is very small, shown by a linear regression slope of  $-4.26 \times 10^{-6}$  ppm s<sup>-1</sup> (Fig. 3c).

### 3.2 CRAM laboratory measurement of cylinder gases

The G2207-*i* analyser performance was evaluated by measuring six gas cylinders with precisely defined O<sub>2</sub> and CO<sub>2</sub> values as measured on a VUV O<sub>2</sub> analyser and Siemens Ultramat 6F NDIR CO<sub>2</sub> analyser (Table 1). The difference between the O<sub>2,NC</sub> values (per meg) as measured by the G2207-*i* and the declared values from the VUV are shown in Table 2 for runs both with and without the RT interpolation applied. This procedure was carried out twice, referred to as “Run 1” and “Run 2” in Table 2.

For both sets of runs without the application of the RT interpolation the difference between the VUV declared value and that measured by the G2207-*i* is very large and far outside of an acceptable range (Table 2), with an average difference from the declared values for all cylinders of  $22.0 \pm 10.3$  per meg. For all cylinders, except for cylinder 5 and 6, a large improvement in the difference is seen after the application of the RT correction. Due to the large differences between the declared and measured values without the RT correction applied, only the results with the RT correction will be discussed hereafter.

Cylinders 5 and 6 contain O<sub>2</sub> values far higher than that found in ambient air (411.7 and 434.6 per meg, respectively) and outside of the range spanned by the WSSes used for calibration. For these two cylinders, the difference between the declared value and that measured by the G2207-*i* is far larger than the other cylinders and also more variable between the two runs, with a standard deviation of the absolute values between the two runs of  $\pm 14.9$  and  $\pm 19.5$  per meg, respectively (Table 2). Berhanu et al. (2019) found that the accuracy of the G2207-*i* was reduced when the CO<sub>2</sub> mole fraction was much higher than that of ambient air but did not observe the same reduction in accuracy with high O<sub>2</sub> mole fractions. Ignoring the two cylinders with positive O<sub>2</sub>, the average absolute difference for the remaining four unknown cylinders and the declared values over the two runs is  $3.4 \pm 2.5$  per meg, this is slightly greater than the WMO compatibility goal of  $\pm 2$  per meg but does fall within the extended goal of  $\pm 10$  per meg and is similar to what can be achieved with an Oxzilla II (Pickers et al., 2017). There is also no correlation between the accuracy and the declared O<sub>2</sub> value when excluding the two cylinders with positive O<sub>2</sub> ( $R^2 = 0.07$  for run 1,  $R^2 = 0.53$  for run 2).

Although the accuracy of the O<sub>2</sub> values measured by the G2207-*i* for these cylinders is variable, particularly for the cylinders with high O<sub>2</sub>, the standard deviation of the 2 min data points used to calculate the final cylinder O<sub>2</sub> value as defined by the G2207-*i* within each run is more consistent. However, the repeatability, used as a proxy for precision, and defined here as the  $\pm 1\sigma$  standard deviation of the average of the two measurements of each cylinder, is variable. For the two cylinders with high O<sub>2</sub> (cylinders 5 and 6) the repeatability is more than 5 times greater than the WMO extended repeatability goal of  $\pm 5$  per meg. For the remaining

**Table 2.** The difference between the O<sub>2</sub> value of each cylinder as measured on the G2207-*i* and the VUV analyser (G2207-*i*–VUV) for two runs on the G2207-*i*, each with and without RT correction applied.

Cylinder no.	Declared O <sub>2</sub> (per meg)	Without RT correction			With RT correction		
		Run 1 difference from declared (per meg) <sup>a</sup>	Run 2 difference from declared (per meg) <sup>a</sup>	Mean of absolute differences of both runs (per meg) <sup>b</sup>	Run 1 difference from declared (per meg) <sup>a</sup>	Run 2 difference from declared (per meg) <sup>a</sup>	Mean of absolute differences of both runs (per meg) <sup>b</sup>
1	−914.8 ± 0.7	9.9 ± 8.4	21.4 ± 8.2	15.7 ± 8.1	0.4 ± 8.5	2.4 ± 8.1	1.4 ± 1.4
2	−880.5 ± 0.9	13.7 ± 8.7	26.5 ± 8.3	20.1 ± 9.1	6.1 ± 8.4	7.6 ± 8.2	6.9 ± 1.1
3	−582.0 ± 1.0	8.1 ± 8.5	22.4 ± 11.3	15.3 ± 10.1	0.7 ± 8.0	3.1 ± 11.2	1.9 ± 1.7
4	−375.0 ± 1.3	12.4 ± 11.6	18.4 ± 9.5	15.4 ± 4.2	5.8 ± 11.3	−1.1 ± 9.5	3.5 ± 3.3
5	411.7 ± 2.1	44.0 ± 12.6	−3.6 ± 11.5	23.8 ± 28.6	19.0 ± 12.4	−40.1 ± 10.2	29.6 ± 14.9
6	434.6 ± 0.3	44.6 ± 5.4	−39.1 ± 10.2	41.9 ± 3.9	22.2 ± 5.1	−49.8 ± 11.5	36.0 ± 19.5

<sup>a</sup> ±1σ standard deviation of the 12 min G2207-*i* average.<sup>b</sup> ±1σ standard deviation of the average of the run 1 and run 2 G2207-*i*–VUV absolute differences.

four cylinders the repeatability is far lower, with cylinder 1 and cylinder 3 both falling within the extended repeatability goal.

### 3.3 Weybourne Atmospheric Observatory field tests

#### 3.3.1 Partial and no drying of ambient air measurements

The results from no drying and partial drying of the sample air into the G2207-*i* at WAO are displayed in Figs. 4 and 5, respectively. The O<sub>2</sub> mole fractions reported (in ppm) by the G2207-*i* were converted to per meg units using the calibration equations produced through the measurement of the three WSS cylinders every 23 h and the concurrent CO<sub>2</sub> observations from the Ultramat 6E analyser.

During the period where there was no drying of the G2207-*i* air sample there is a significant difference between the O<sub>2</sub> values reported by the Oxzilla (dried air) and the G2207-*i* O<sub>2,NC</sub> values (Fig. 4b). This is to be expected due to the diluting effect of water vapour; however, there is also a significant difference between the Oxzilla O<sub>2</sub> and the G2207-*i* O<sub>2,WC</sub> values. Over the entire no-drying period the average difference between the Oxzilla observations and the G2207-*i* O<sub>2,NC</sub> is −965.4 ± 272.8 per meg. The average difference between the Oxzilla and the G2207-*i* O<sub>2,WC</sub> values is −849.8 ± 31.1 per meg. Although the difference is substantially smaller with the application of the G2207-*i* built-in water correction procedure, it is still unusably large, with no similarity in the Oxzilla and G2207-*i* signals and both the O<sub>2,NC</sub> and O<sub>2,WC</sub> G2207-*i* values correlating with the H<sub>2</sub>O variability (Fig. 6a and b). This demonstrates that the algorithm currently applied for water correction is unsuitable for precise O<sub>2</sub> measurement.

As seen during the no-drying period of the sample air, there is also a significant difference between the reported O<sub>2</sub>

values of the Oxzilla and G2207-*i* under the partial drying regime for both O<sub>2,NC</sub> and O<sub>2,WC</sub> (Fig. 5b). With partial drying, the time series of the difference between the O<sub>2</sub> values of the two analysers is a lot smoother than with no drying. This is due to the fridge trap removing some of the natural variability in the water vapour mole fraction. Over the entire partial drying period the average difference between the Oxzilla observations and the G2207-*i* O<sub>2,NC</sub> is −7144.1 ± 258.6 per meg. The average difference between the Oxzilla and the G2207-*i* O<sub>2,WC</sub> values is −612.7 ± 31.8 per meg. There is a large improvement with the application of the water correction procedure; however, as with the no-drying results, the difference in O<sub>2</sub> values between the Oxzilla and G2207-*i* O<sub>2,WC</sub> are too large to be usable for any application, with the O<sub>2,NC</sub> and O<sub>2,WC</sub> values correlating with the H<sub>2</sub>O variability (Fig. 6c and d).

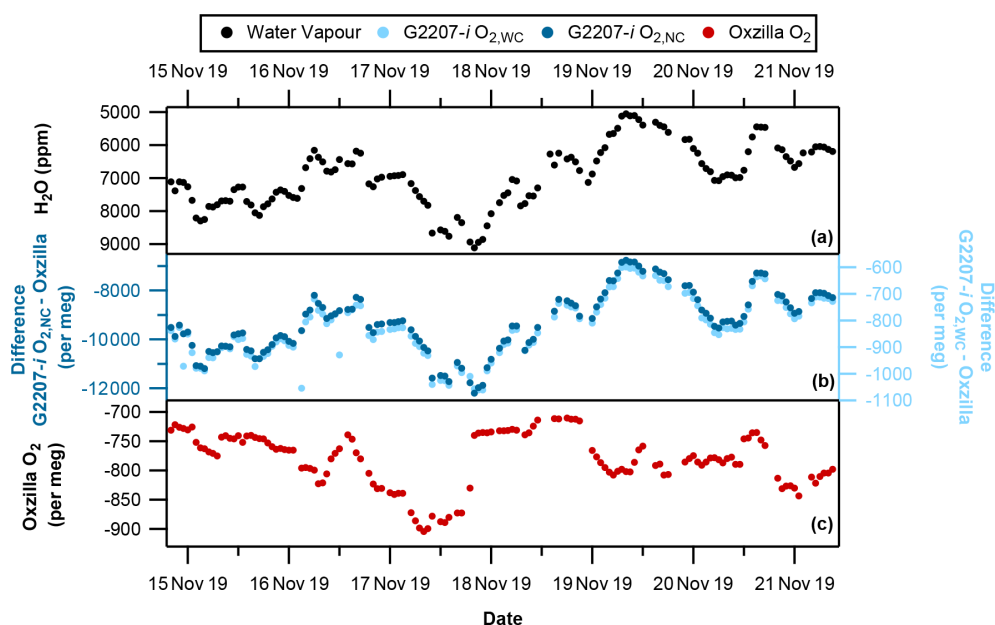
Under both partial-drying and no-drying regimes, the difference between the Oxzilla and G2207-*i* values is strongly correlated with the water vapour mole fraction but decreases with the application of the built-in water correction procedure (Fig. 6). The *R*<sup>2</sup> value decreases from 0.996 to 0.803 for no drying and from 0.967 to 0.301 for partial drying once the water correction has been applied. Given the correlation between the water vapour mole fraction and the O<sub>2,WC</sub> reported by the G2207-*i* these values are not usable without significant improvements to the water correction procedure by Picarro Inc.

Due to the large differences observed between the Oxzilla and G2207-*i* reported O<sub>2</sub> values under no drying and partial drying, no further investigation was undertaken, thus only the fully dried sample air data are considered hereafter.

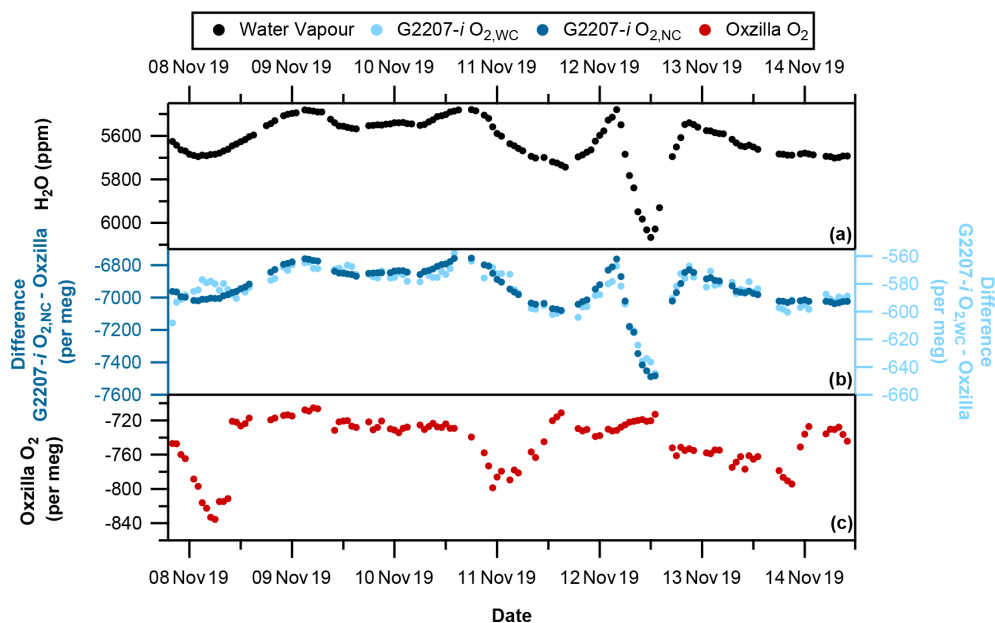
#### 3.3.2 Full drying of ambient air measurements

The results from fully drying the sample air between 24 October and 7 November 2019 are displayed in Fig. 7. The O<sub>2</sub>





**Figure 4.** With no drying of the sample air through the G2207-*i* (a) hourly averaged water vapour, (b) G2207-*i*–Oxzilla difference for O<sub>2,NC</sub> (dark blue) and O<sub>2,WC</sub> (light blue), and (c) Oxzilla O<sub>2</sub>. Note the reversed water vapour axis and different axis scales for O<sub>2,NC</sub> and O<sub>2,WC</sub>.

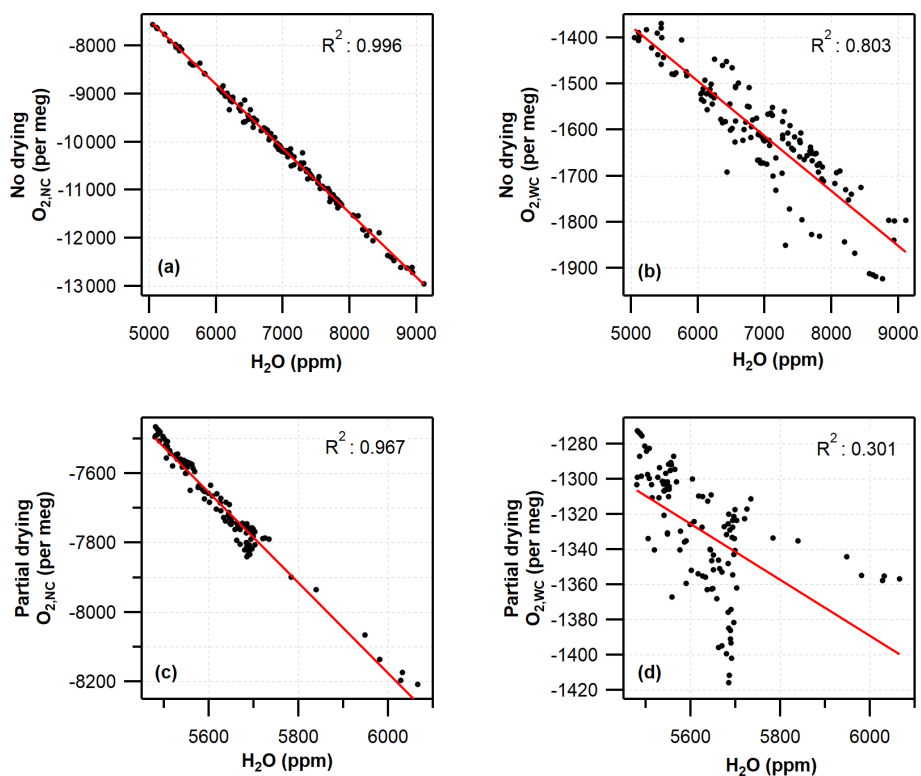


**Figure 5.** With partial drying of the sample air through the G2207-*i* (a) Hourly averaged water vapour, (b) G2207-*i*–Oxzilla difference for O<sub>2,NC</sub> (dark blue) and O<sub>2,WC</sub> (light blue), and (c) Oxzilla O<sub>2</sub> (Oxzilla sample air is fully dried). Note the reversed water vapour axis and different axis scales for O<sub>2,NC</sub> and O<sub>2,WC</sub>. The spike in water vapour on 12 November 2019 is due to a temporary increase in the temperature of the fridge.

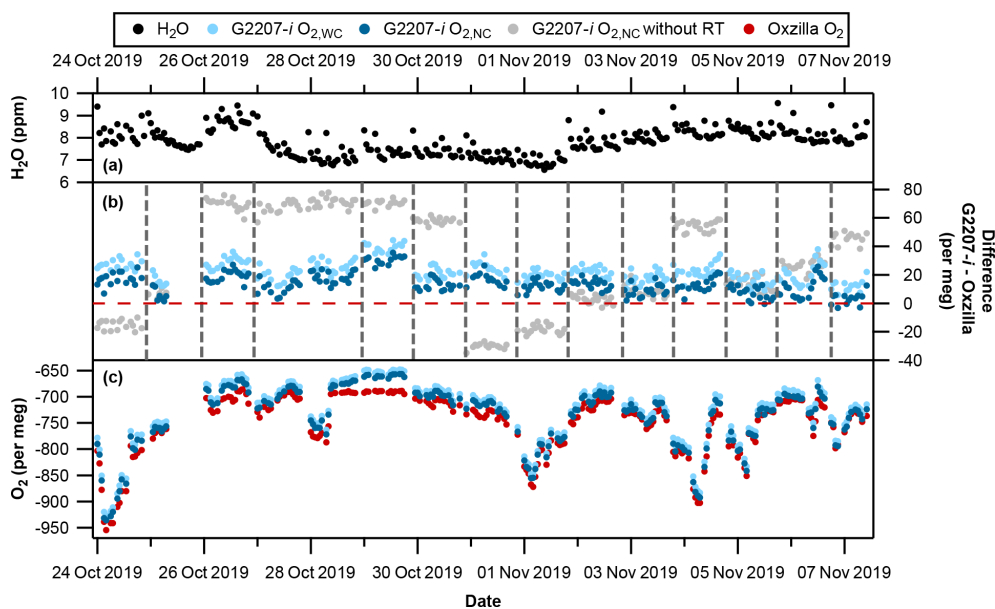
mole fractions reported in ppm units by the G2207-*i* were converted to per meg units using the calibration equations produced through the measurement of the three WSS cylinders every 23 h, and the concurrent CO<sub>2</sub> observations.

There is a greater difference between the Oxzilla and G2207-*i* O<sub>2,WC</sub> values than the O<sub>2,NC</sub> values, with an av-

erage difference over the entire full-drying period of  $22.6 \pm 7.4$  per meg compared to  $13.6 \pm 7.5$  per meg, respectively. This may be due to overcorrection of the O<sub>2,NC</sub> values as the water vapour mole fraction is below the G2207-*i*'s lower detection limit and precision, i.e. the G2207-*i* is reporting H<sub>2</sub>O mole fractions of approximately 7 ppm (Fig. 7a) (with



**Figure 6.** Correlation between water vapour mole fraction and hourly averaged G2207-*i* O<sub>2</sub> for (a) no-drying O<sub>2,NC</sub>, (b) no-drying O<sub>2,WC</sub>, (c) partial-drying O<sub>2,NC</sub>, and (d) partial-drying O<sub>2,WC</sub>. Red lines show linear regression.



**Figure 7.** Time series with full drying of the air sample. (a) Hourly averaged water vapour; spikes are due to equilibration after valve switching from cylinder air to sample air. (b) G2207-*i*-Oxzilla difference for O<sub>2,WC</sub> (light blue), O<sub>2,NC</sub> (dark blue), and O<sub>2,NC</sub> without the RT interpolation applied (grey). Vertical dashed lines indicate a full three-gas WSS calibration on the G2207-*i*, and the red horizontal line indicates zero difference from the Oxzilla. (c) Hourly averaged Oxzilla O<sub>2</sub> (red), O<sub>2,WC</sub> (light blue), and O<sub>2,NC</sub> (dark blue). Note that there was no WSS calibration on 27 October 2019 due to a macro error which prevented valve switching to calibration gases; therefore, the calibration from 26 October 2019 was applied for 46 h.

frequent spikes due to equilibration after switching of V1 (Fig. 1) from cylinder to sample air); however, when the air sample is fully dried by passing through the chiller and fridge trap, the water vapour is reduced to below 1 ppm. This overestimated water correction whilst sampling fully dried air was also found by Berhanu et al. (2019). We therefore only refer to the O<sub>2,NC</sub> values, which we believe to be more accurate, in the analysis from now onwards.

The large jumps in the G2207-*i* O<sub>2,NC</sub> values following WSS calibrations (see Fig. 7b, grey points) are caused by a drift in the analyser's baseline, which only becomes applied to the data after each calibration. These jumps were reduced through the application of the 5 h RT interpolation procedure (see Fig. 7b, blue points), which constrained the baseline drift (refer to Sect. 2.3.2). After the application of the RT interpolation the jumps between WSS calibrations were vastly reduced (see Fig. 7).

### 3.3.3 Repeatability and compatibility

The repeatability and compatibility of the analyser were evaluated through the running of a TT every 7 h during the full-drying period using O<sub>2,NC</sub> values, the results of which are presented in Fig. 8 and Table 3. For O<sub>2</sub> the WMO repeatability goal is  $\pm 1$  per meg (with an extended goal of  $\pm 5$  per meg) and the compatibility goal is  $\pm 2$  per meg (with an extended goal of  $\pm 10$  per meg; indicated by the dashed lines in Fig. 8; Crotwell et al., 2019).

The repeatability is determined from the mean  $\pm 1\sigma$  standard deviations of the average of two consecutive measurements of the TT. For the G2207-*i* this is equal to  $\pm 5.7 \pm 5.6$  per meg, compared to  $\pm 2.2 \pm 2.0$  per meg on the Oxzilla. Prior to applying the RT interpolation to the G2207-*i* data, the repeatability of the G2207-*i* was  $\pm 11.9 \pm 13.8$  per meg, twice as bad as after the RT application; this is because after the RT interpolation was applied the large jumps in the TT value after a WSS calibration were removed. In the context of the WMO repeatability goals, neither the Oxzilla nor the G2207-*i* meet the goal of  $\pm 1$  per meg. For O<sub>2</sub>, the WMO goals are very ambitious and not currently achievable by the O<sub>2</sub> measurement community; hence, the “extended” O<sub>2</sub> repeatability goal of  $\pm 5$  per meg (Crotwell et al., 2019). The Oxzilla TT results lie within this extended goal; however, the G2207-*i* does not, even after the application of the RT interpolation.

The compatibility of the analyser, which is used here as a proxy for accuracy, is determined by calculating the mean difference between the TT O<sub>2</sub> as measured by the G2207-*i* and the VUV declared value ( $-718$  per meg). The mean absolute difference from the declared value on the VUV for the Oxzilla is  $3.0 \pm 2.6$  per meg, this is well within the extended WMO compatibility goal of  $\pm 10$  per meg and is quite close to the more stringent goal of  $\pm 2$  per meg. The compatibility of the G2207-*i* prior to the application of the RT interpolation is  $22.9 \pm 34.1$  per meg, which is far greater than even the ex-

tended compatibility goal of  $\pm 10$  per meg. After the application of the RT interpolation the compatibility of the G2207-*i* O<sub>2,NC</sub> is  $10.0 \pm 6.7$  per meg, although this is not within the WMO compatibility goal, it is just within the extended goal, which is deemed suitable for some applications in specific circumstances, such as where the signals are relatively very large so that reduced repeatability and compatibility does not preclude useful information from the measurements.

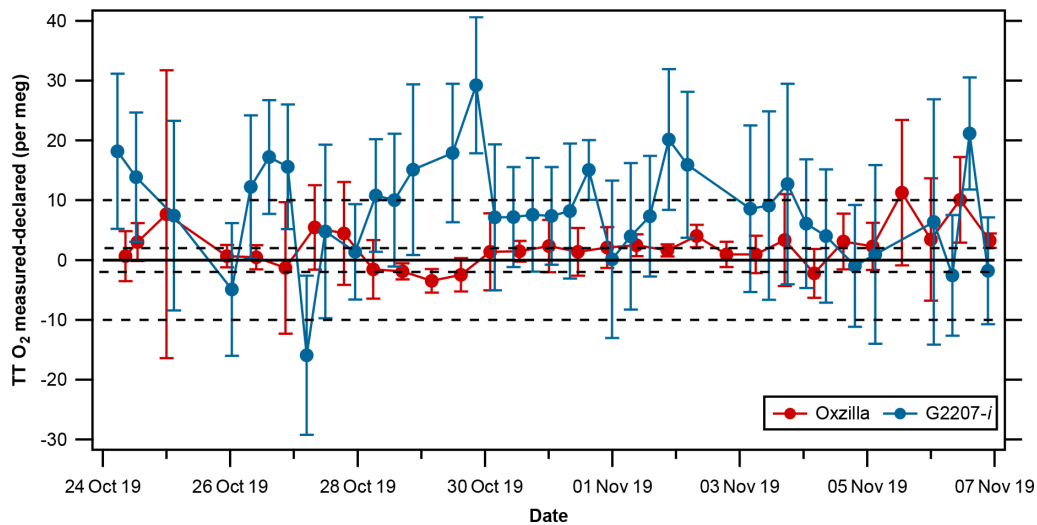
The compatibility and repeatability of the G2207-*i* measurements were vastly improved after the application of a 5-hourly RT; however, if one ignores the TT results immediately after a new WSS calibration (i.e. after the large jumps when the RT was not applied), the repeatability without the RT interpolation is  $5.2 \pm 4.5$  per meg, improving to  $4.3 \pm 4.6$  per meg when the RT is applied. This is because the RT corrected for baseline drift between WSS calibrations, but it does not correct for drift within the calibration period. However, as the TT results are imprecise (as illustrated by the large error bars in Fig. 8), even if any baseline drift within a calibration period were corrected for, there would likely be little improvement in the final TT results as the noise in the RT-corrected TT values is primarily caused by imprecision rather than baseline drift.

### 3.4 Applications of the G2207-*i* O<sub>2</sub> measurements in the calculation of fossil fuel CO<sub>2</sub>

In order to further assess the G2207-*i*'s performance in real-world applications the fully dried, RT corrected, O<sub>2,NC</sub> observations from WAO were used to isolate the fossil fuel component of the concurrent CO<sub>2</sub> observations and then compared to the ffCO<sub>2</sub> values calculated from the Oxzilla O<sub>2</sub> observations following the APO methodology outlined in Pickers et al. (2022). The resultant ffCO<sub>2</sub> values calculated from each analyser are displayed in Fig. 9.

The measurement uncertainty was calculated as the average hourly SD on 30 October 2019, this date was chosen as it was a particularly stable period with little variation in the TT results for both analysers (Fig. 8); the resultant uncertainty for the G2207-*i* is  $\pm 11.2$  per meg compared to  $\pm 4.9$  per meg for the Oxzilla. The uncertainty in the baseline determination ( $\pm 28\%$ ) and the emission ratio uncertainty ( $\pm 22\%$ ) are significantly larger than these measurement uncertainties (Pickers et al., 2022), but as these are the same for both analysers the additional measurement uncertainty for the G2207-*i* caused by analyser noise increases the uncertainty of the calculated ffCO<sub>2</sub> values. The average final calculated uncertainty on the ffCO<sub>2</sub> values calculated from the Oxzilla measurements is  $\pm 5.8$  ppm, compared to  $\pm 13.0$  ppm on the G2207-*i*.

The average ffCO<sub>2</sub> value over the entire full-drying period for the Oxzilla is 5.1 ppm, compared to 7.9 ppm on the G2207-*i* (Table 4); the calculated ffCO<sub>2</sub> from the G2207-*i* is higher than that of the Oxzilla 73% of the time. This difference is predominantly due to the higher O<sub>2</sub> values re-



**Figure 8.** TT differences from declared values (measured–declared) ( $\pm 1\sigma$  standard deviation) for the Oxzilla (red) and G2207-*i* O<sub>2,NC</sub> (blue). The solid line indicates zero difference from the declared O<sub>2</sub> value of the TT, and the dashed lines indicate the WMO compatibility goal of  $\pm 2$  per meg and the extended goal of  $\pm 10$  per meg.

**Table 3.** Repeatability and compatibility goals and achievements for each analyser.

	Repeatability (per meg) <sup>a</sup>	Compatibility (per meg) <sup>b</sup>
WMO compatibility goal	$\pm 1$ ( $\pm 5$ ) <sup>c</sup>	$\pm 2$ ( $\pm 10$ ) <sup>c</sup>
Oxzilla	$\pm 2.2 \pm 2.0$	$\pm 3.0 \pm 2.6$
G2207- <i>i</i> O <sub>2,NC</sub> without RT interpolation	$\pm 11.9 \pm 13.8$	$\pm 22.9 \pm 34.1$
G2207- <i>i</i> O <sub>2,NC</sub> with RT interpolation	$\pm 5.7 \pm 5.6$	$\pm 10.0 \pm 6.7$

<sup>a</sup> Values are calculated using the method in Kozlova and Manning (2009) and Pickers et al. (2017). Mean  $\pm 1\sigma$  standard deviations of the average of two consecutive measurements of the TT, determined from 30 TT measurements for the Oxzilla and 37 TT measurements for the G2207-*i*, where one run is the average of 12 min of data. Uncertainties are given on these mean standard deviations, illustrating that the analytical repeatability is variable over time.

<sup>b</sup> Mean differences between the measured TT O<sub>2</sub>/N<sub>2</sub> ratio and the declared values determined on the VUV analyser against primary calibration standards on the SIO O<sub>2</sub> scale.

<sup>c</sup> WMO repeatability and compatibility goals, where the repeatability of a measurement should be at most half of the value of the compatibility goal. For O<sub>2</sub>, the WMO goals are very ambitious and not currently achievable by the O<sub>2</sub> measurement community; hence, the “extended” O<sub>2</sub> goals, which are suitable for some O<sub>2</sub> applications, shown in parentheses.

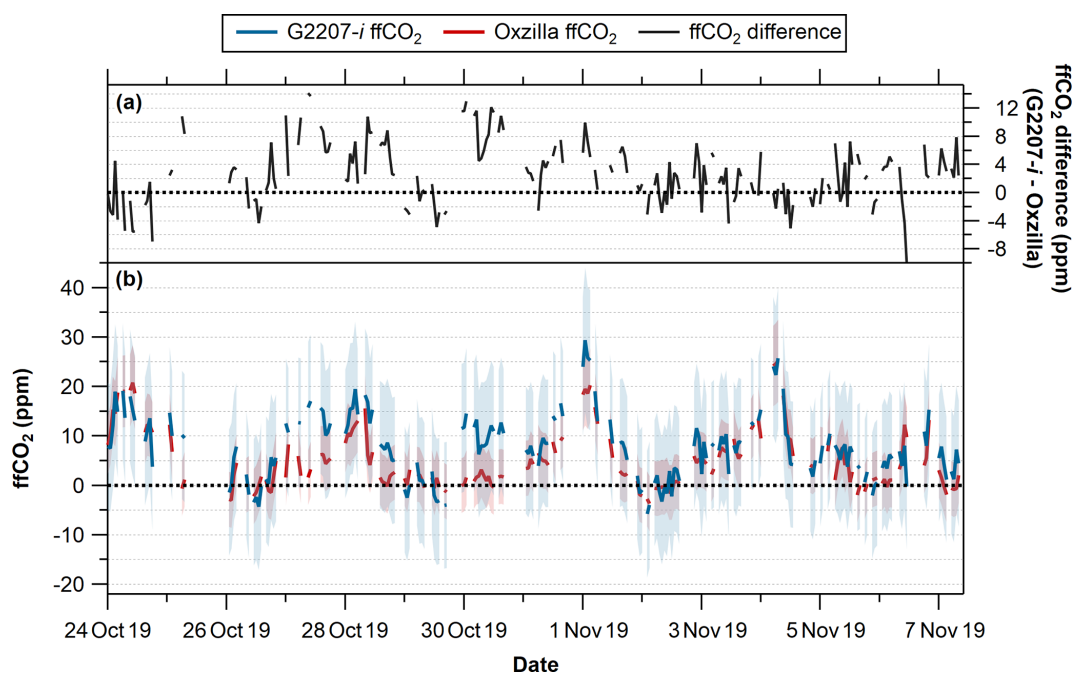
ported by the G2207-*i* as discussed in Sect. 3.3.2; some of this difference also comes from the jumps in the G2207-*i* O<sub>2</sub> values, which means that the calculated baselines used for each analyser follow different trends. For example, on the 27 and 30 October 2019 the largest difference between the calculated ffCO<sub>2</sub> values is observed (Fig. 9); on both of these dates there is a large jump in O<sub>2</sub> values from the previous day measured by the G2207-*i* following a WSS calibration (Fig. 7). Although the O<sub>2</sub> difference between the two analysers on these days is low, there was a large difference the preceding day, the days with the larger difference (due to a higher O<sub>2</sub> value reported by the G2207-*i*) in observed values pull the baseline to become more positive, thus making the difference between the ffCO<sub>2</sub> calculated from the two analysers larger on days where the observed O<sub>2</sub> difference is smaller.

Although the G2207-*i* calculated ffCO<sub>2</sub> values that are often higher than those from the Oxzilla, they still follow the

**Table 4.** Comparison of ffCO<sub>2</sub> values calculated from the Oxzilla and G2207-*i* O<sub>2</sub> measurements. Average values are given with  $\pm 1\sigma$  standard deviation.

	Oxzilla ffCO <sub>2</sub> (ppm)	G2207- <i>i</i> ffCO <sub>2</sub> (ppm)
Average	$5.1 \pm 5.9$	$7.9 \pm 6.6$
Maximum	25.2	29.4
Minimum	−3.7	−6.5

same trend (with some jumps in the G2207-*i* values); however, the maximum and minimum values occur at different times. The differences in ffCO<sub>2</sub> calculated from the G2207-*i* and the Oxzilla would become problematic if using the G2207-*i* analyser for top-down ffCO<sub>2</sub> quantification on an hourly basis.



**Figure 9.** (a) Difference between the ffCO<sub>2</sub> calculated using O<sub>2</sub> from the G2207-*i* and the Oxzilla (G2207-*i* – Oxzilla). (b) Calculated ffCO<sub>2</sub> using O<sub>2</sub> from the G2207-*i* (blue) and the Oxzilla (red); blue and red shaded areas indicate uncertainty of the calculated ffCO<sub>2</sub>. Dashed black lines indicate 0 ppm. Negative ffCO<sub>2</sub> values occur when the O<sub>2</sub> observations are above (more positive than) the calculated O<sub>2</sub> baseline. Note that gaps are due to the threshold requirement of a minimum of 20 min of data for hourly averages.

#### 4 Conclusions

The performance of the Picarro G2207-*i* under both laboratory and field conditions has been thoroughly evaluated. When running a cylinder on the G2207-*i* over 24 h in the laboratory, we observed a large amount of noise in the raw 1 s data, resulting in a large standard deviation in averaged data. This standard deviation is reduced over longer averaging times. During the laboratory measurement of cylinder gases with declared O<sub>2</sub> values, the G2207-*i* performed within the WMO extended compatibility goal of  $\pm 10$  per meg when measuring cylinders with a negative O<sub>2</sub> per meg value. When measuring cylinders with a positive O<sub>2</sub> value, the precision and accuracy of the result worsened, thus the G2207-*i* is not recommended for use in this range.

When sampling ambient air, we found that the G2207-*i*'s built-in water correction does not, at present, sufficiently correct for the influence of water vapour even when the sample air is partially dried, and we therefore recommend full drying ( $< 1$  ppm H<sub>2</sub>O) of air samples. When sampling fully dried air, large step-changes in the reported O<sub>2</sub> values from the G2207-*i* were observed after each WSS calibration; the addition of a RT every 5 h vastly reduced these jumps; however, they were still observable. When the RT interpolation was applied, the repeatability of the G2207-*i* was  $\pm 5.7 \pm 5.6$  per meg, falling just outside of the WMO extended goal of  $\pm 5$  per meg; it is possible that with a more frequent RT interpolation this repeatability will improve. The compatibility

was  $\pm 10 \pm 6.7$  per meg, falling within the WMO extended compatibility goal for O<sub>2</sub> of  $\pm 10$  per meg. In the future, investigation into whether increasing the frequency of the running of an RT to reduce jumps in the observed O<sub>2</sub> values after a WSS calibration may improve both the repeatability and compatibility of the analyser. A key benefit of CRDS analysers is that they do not require drying of the air sample; however, this is not currently the case with the G2207-*i* for O<sub>2</sub> measurements.

*Data availability.* The G2207-*i* data from the WAO and CRAM Lab tests are available at: <https://doi.org/10.5281/zenodo.6802657> (Fleming et al., 2022). The WAO in situ datasets are available at the CEDA data archives, for O<sub>2</sub>: <https://catalogue.ceda.ac.uk/uuid/b3f9714c956f428a840211e0184e23eb> (last access: 1 July 2022; Forster, 2012b), and for CO<sub>2</sub>: <https://catalogue.ceda.ac.uk/uuid/87fc265aab6b4aeb961e62da2cd6ca91> (last access: 1 July 2022; Forster, 2012a).

*Author contributions.* LSF, ACM, and PAP developed the measurement methodology, and the measurements were conducted by LSF at UEA and WAO. AJE developed the software used to run the analyser. Investigation and visualisation were completed by LSF. Writing was carried out by LSF. Reviewing and editing were done by LSF, ACM, PAP, and GLF.

*Competing interests.* The contact author has declared that none of the authors has any competing interests.

*Disclaimer.* Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

*Acknowledgements.* We are very grateful to Marica Hewitt, Nick Griffin, and Dave Blomfield (UEA) for supporting the WAO measurements. We are also very grateful to Gregor Lucic and Magdalena Hofmann at Picarro Inc. for the loaning of the G2207-*i* analyser and their feedback on a draft manuscript.

*Financial support.* Leigh S. Fleming was supported by the UK Natural Environment Research Council (NERC) "EnvEast" Doctoral Training Partnership (grant no. NE/L002582/1). The WAO atmospheric O<sub>2</sub> and CO<sub>2</sub> measurements are supported by the Atmospheric Measurement and Observation Facility (AMOF) of the National Centre for Atmospheric Science (NCAS), in addition to NERC research grant nos. NE/R011532/1 and NE/S004521/1.

*Review statement.* This paper was edited by Huilin Chen and reviewed by two anonymous referees.

## References

- Battle, M. O., Munger, J. W., Conley, M., Sofen, E., Perry, R., Hart, R., Davis, Z., Scheckman, J., Woogerd, J., Graeter, K., Seekins, S., David, S., and Carpenter, J.: Atmospheric measurements of the terrestrial O<sub>2</sub>:CO<sub>2</sub> exchange ratio of a midlatitude forest, *Atmos. Chem. Phys.*, 19, 8687–8701, <https://doi.org/10.5194/acp-19-8687-2019>, 2019.
- Bender, M. L., Tans, P. P., Ellis, J. T., Orchardo, J., and Habfast, K.: A high precision isotope ratio mass spectrometry method for measuring the O<sub>2</sub>/N<sub>2</sub> ratio of air, *Geochim. Cosmochim. Acta*, 58, 4751–4758, [https://doi.org/10.1016/0016-7037\(94\)90205-4](https://doi.org/10.1016/0016-7037(94)90205-4), 1994.
- Berhanu, T. A., Hoffnagle, J., Rella, C., Kimhak, D., Nyfeler, P., and Leuenberger, M.: High-precision atmospheric oxygen measurement comparisons between a newly built CRDS analyzer and existing measurement techniques, *Atmos. Meas. Tech.*, 12, 6803–6826, <https://doi.org/10.5194/amt-12-6803-2019>, 2019.
- Blaine, T. W., Keeling, R. F., and Paplawsky, W. J.: An improved inlet for precisely measuring the atmospheric Ar/N<sub>2</sub> ratio, *Atmos. Chem. Phys.*, 6, 1181–1184, <https://doi.org/10.5194/acp-6-1181-2006>, 2006.
- Chen, H., Winderlich, J., Gerbig, C., Hofer, A., Rella, C. W., Crosson, E. R., Van Pelt, A. D., Steinbach, J., Kolle, O., Beck, V., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Santoni, G. W., and Wofsy, S. C.: High-accuracy continuous airborne measurements of greenhouse gases (CO<sub>2</sub> and CH<sub>4</sub>) using the cavity ring-down spectroscopy (CRDS) technique, *Atmos. Meas. Tech.*, 3, 375–386, <https://doi.org/10.5194/amt-3-375-2010>, 2010.
- Crotwell, A., Lee, H., and Steinbacher, M.: Report of the 20th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Measurement Techniques (GGMT-2019), World Meteorological Organization Global Atmosphere Watch, Jeju Island, South Korea Report Series, GAW Report No. 255, [https://library.wmo.int/doc\\_num.php?explnum\\_id=10353](https://library.wmo.int/doc_num.php?explnum_id=10353) (last access: 1 July 2022), 2019.
- Dlugokencky, E. J. and Tans, P. P.: Trends in Atmospheric Carbon Dioxide. National Oceanic and Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL), [https://gml.noaa.gov/ccgg/trends/gl\\_gr.html](https://gml.noaa.gov/ccgg/trends/gl_gr.html), last access: 17 October 2022.
- Fleming, L. S., Manning, A. C., Pickers, P. A., Forster, G. L., and Etchells, A. J.: Datasets for "Evaluating the performance of a Picarro G2207-*i* analyser for high-precision atmospheric O<sub>2</sub> measurements", Zenodo [data set], <https://doi.org/10.5281/zenodo.6802657>, 2022.
- Forster, G.: Weybourne Atmospheric Observatory: Longterm measurements of Atmospheric Carbon Dioxide, NCAS British Atmospheric Data Centre [data set], <https://catalogue.ceda.ac.uk/uuid/87fc265aab6b4aeb961e62da2cd6ca91>, 2012a.
- Forster, G.: Weybourne Atmospheric Observatory: Long term measurements of atmospheric O<sub>2</sub>, NCAS British Atmospheric Data Centre [data set], <https://catalogue.ceda.ac.uk/uuid/b3f9714c956f428a840211e0184e23eb> (last access: 1 July 2022), 2012b.
- ICOS-RI: ICOS Atmosphere Station Specifications V2.0, ICOS ERIC, <https://doi.org/10.18160/GK28-2188>, 2020.
- Keeling, R. F.: Measuring correlations between atmospheric oxygen and carbon-dioxide mole fractions – a preliminary-study in urban air, *J. Atmos. Chem.*, 7, 153–176, <https://doi.org/10.1007/bf00048044>, 1988a.
- Keeling, R. F.: Development of an Interferometric Oxygen Analyzer for Precise Measurement of the Atmospheric O<sub>2</sub> Mole Fraction, PhD thesis, Division of Applied Sciences, Harvard University, Cambridge, Massachusetts, USA, 178 pp., <https://scripps2.ucsd.edu/scientific-literature.html> (last access: 1 July 2022), 1988b.
- Keeling, R. F. and Manning, A. C.: 5.15 – Studies of Recent Changes in Atmospheric O<sub>2</sub> Content, in: *Treatise on Geochemistry*, 2nd edn., edited by: Holland, H. D. and Turekian, K. K., Elsevier, Oxford, 385–404, <https://doi.org/10.1016/B978-0-08-095975-7.00420-4>, 2014.
- Keeling, R. F. and Shertz, S. R.: Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle, *Nature*, 358, 723–727, <https://doi.org/10.1038/358723a0>, 1992.
- Keeling, R. F., Manning, A. C., McEvoy, E. M., and Shertz, S. R.: Methods for measuring changes in atmospheric O<sub>2</sub> concentration and their application in southern hemisphere air, *J. Geophys. Res.-Atmos.*, 103, 3381–3397, <https://doi.org/10.1029/97JD02537>, 1998.
- Keeling, R. F., Manning, A. C., Paplawsky, W. J., and Cox, A. C.: On the long-term stability of reference gases for atmospheric O<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub> measurements, *Tellus B*, 59, 3–14, <https://doi.org/10.1111/j.1600-0889.2006.00228.x>, 2007.
- Kozlova, E. A. and Manning, A. C.: Methodology and calibration for continuous measurements of biogeochemical trace gas and O<sub>2</sub> concentrations from a 300-m tall tower in central Siberia,

- Atmos. Meas. Tech., 2, 205–220, <https://doi.org/10.5194/amt-2-205-2009>, 2009.
- Manning, A. C.: Temporal variability of atmospheric oxygen from both continuous measurements and a flask sampling network: Tools for studying the global carbon cycle, PhD thesis, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA, 202 pp., <https://cramlab.uea.ac.uk/Publications.php> (last access: 1 July 2022), 2001.
- Manning, A. C., Keeling, R. F., and Severinghaus, J. P.: Precise atmospheric oxygen measurements with a paramagnetic oxygen analyzer, *Glob. Biogeochem. Cy.*, 13, 1107–1115, <https://doi.org/10.1029/1999GB900054>, 1999.
- Pickers, P. A.: New applications of continuous atmospheric O<sub>2</sub> measurements: Meridional transects across the Atlantic Ocean, and improved quantification of fossil fuel-derived CO<sub>2</sub>, PhD thesis, School of Environmental Sciences, University of East Anglia, Norwich, UK, 262 pp., <https://cramlab.uea.ac.uk/Publications.php> (last access: 1 July 2022), 2016.
- Pickers, P. A., Manning, A. C., Sturges, W. T., Le Quéré, C., Mikaloff Fletcher, S. E., Wilson, P. A., and Etchells, A. J.: In situ measurements of atmospheric O<sub>2</sub> and CO<sub>2</sub> reveal an unexpected O<sub>2</sub> signal over the tropical Atlantic Ocean, *Glob. Biogeochem. Cy.*, 31, 1289–1305, <https://doi.org/10.1002/2017GB005631>, 2017.
- Pickers, P. A., Manning, A. C., Quéré, C. L., Forster, G. L., Luijkx, I. T., Gerbig, C., Fleming, L. S., and Sturges, W. T.: Novel quantification of regional fossil fuel CO<sub>2</sub> reductions during COVID-19 lockdowns using atmospheric oxygen measurements, *Sci. Adv.*, 8, eab19250, <https://doi.org/10.1126/sciadv.ab19250>, 2022.
- Rella, C. W., Chen, H., Andrews, A. E., Filges, A., Gerbig, C., Hatakka, J., Karion, A., Miles, N. L., Richardson, S. J., Steinbacher, M., Sweeney, C., Wastine, B., and Zellweger, C.: High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air, *Atmos. Meas. Tech.*, 6, 837–860, <https://doi.org/10.5194/amt-6-837-2013>, 2013.
- Resplandy, L., Keeling, R. F., Eddebbar, Y., Brooks, M., Wang, R., Bopp, L., Long, M. C., Dunne, J. P., Koeve, W., and Oschlies, A.: Quantification of ocean heat uptake from changes in atmospheric O<sub>2</sub> and CO<sub>2</sub> composition, *Sci. Rep.*, 9, 20244, <https://doi.org/10.1038/s41598-019-56490-z>, 2019.
- Ruckstuhl, A. F., Henne, S., Reimann, S., Steinbacher, M., Vollmer, M. K., O'Doherty, S., Buchmann, B., and Hueglin, C.: Robust extraction of baseline signal of atmospheric trace species using local regression, *Atmos. Meas. Tech.*, 5, 2613–2624, <https://doi.org/10.5194/amt-5-2613-2012>, 2012.
- Severinghaus, J. P.: Studies of the terrestrial O<sub>2</sub> and carbon cycles in sand dune gases and in Biosphere 2, PhD thesis, Graduate School of Arts and Sciences, Columbia University, USA, 159 pp., <https://doi.org/10.2172/477735>, 1995.
- Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., and Mukai, H.: The CO<sub>2</sub> release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, *Atmos. Chem. Phys.*, 11, 6855–6870, <https://doi.org/10.5194/acp-11-6855-2011>, 2011.
- Stephens, B. B., Keeling, R. F., Heimann, M., Six, K. D., Murnane, R., and Caldeira, K.: Testing global ocean carbon cycle models using measurements of atmospheric O<sub>2</sub> and CO<sub>2</sub> concentration, *Glob. Biogeochem. Cy.*, 12, 213–230, <https://doi.org/10.1029/97GB03500>, 1998.
- Stephens, B. B., Bakwin, P. S., Tans, P. P., Teclaw, R. M., and Baumann, D. D.: Application of a differential fuel-cell analyzer for measuring atmospheric oxygen variations, *J. Atmos. Ocean. Technol.*, 24, 82–94, <https://doi.org/10.1175/JTECH1959.1>, 2007.
- Stephens, B. B., Keeling, R. F., and Paplawsky, W. J.: Shipboard measurements of atmospheric oxygen using a vacuum-ultraviolet absorption technique, *Tellus B*, 55, 857–878, <https://doi.org/10.3402/tellusb.v55i4.16386>, 2011.
- Tohjima, Y.: Method for measuring changes in the atmospheric O<sub>2</sub>/N<sub>2</sub> ratio by a gas chromatograph equipped with a thermal conductivity detector, *J. Geophys. Res.-Atmos.*, 105, 14575–14584, <https://doi.org/10.1029/2000JD900057>, 2000.
- Tohjima, Y., Machida, T., Watai, T., Akama, I., Amari, T., and Moriwaki, Y.: Preparation of gravimetric standards for measurements of atmospheric oxygen and reevaluation of atmospheric oxygen concentration, *J. Geophys. Res.-Atmos.*, 110, D11302, <https://doi.org/10.1029/2004JD005595>, 2005.
- Tohjima, Y., Mukai, H., Machida, T., Hoshina, Y., and Nakaoka, S.-I.: Global carbon budgets estimated from atmospheric O<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub> observations in the western Pacific region over a 15-year period, *Atmos. Chem. Phys.*, 19, 9269–9285, <https://doi.org/10.5194/acp-19-9269-2019>, 2019.
- Werle, P., Mücke, R., and Slemr, F.: The limits of signal averaging in atmospheric trace-gas monitoring by tunable diode-laser absorption spectroscopy (TDLAS), *Appl. Phys. B*, 57, 131–139, <https://doi.org/10.1007/BF00425997>, 1993.
- Wilson, P. A.: Insight into the Carbon Cycle from Continuous Measurements of Oxygen and Carbon Dioxide at Weybourne Atmospheric Observatory, UK, PhD thesis, School of Environmental Sciences, University of East Anglia, Norwich, UK, 155 pp., <https://cramlab.uea.ac.uk/Publications.php> (last access: 1 July 2022), 2013.
- Zhao, C. L. and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, *J. Geophys. Res.-Atmos.*, 111, D08S09, <https://doi.org/10.1029/2005JD006003>, 2006.