1 Development of an SI-traceable transmission curve reference material

2 to improve comparability of proton transfer reaction mass

3 spectrometry measurements

4 David R. Worton[1], Sergi Moreno[1,#], Kieran O'Daly[1], Rupert Holzinger[2]

5 [1] National Physical Laboratory, Hampton Road, Teddington, TW11 0LW, United Kingdom

6 [2] Institute for Marine and Atmospheric Research, IMAU, Utrecht University, Utrecht, the Netherlands

7 *Correspondence to*: David R. Worton (dave.worton@npl.co.uk)

8 [#] Now at World Meteorological Organization (WMO), 7bis Avenue de la Paix, C.P. 2300

9 CH-1211, Geneva 2, Switzerland

11 **Abstract.** Since its inception more than two decades ago proton-transfer-reaction mass-spectrometry (PTR-MS) has

12 established itself as a powerful technique for the measurements of volatile organic compounds (VOCs) providing high time

13 resolution, high sensitivity measurements without any sample pre-treatment. As this technology has matured and its application

14 become more widespread there is a growing need for accurate and traceable calibration to ensure measurement comparability.

15 As a result of the large number of VOCs detectable with PTR-MS it is impractical to have a calibration standard or standards

16 that cover all observable compounds. However, recent work has demonstrated that quantitative measurements of uncalibrated

17 compounds are possible provided that the transmission curve is sufficiently constrained. To enable this, a novel multi-

18 component gas reference material containing 20 compounds that spans the molecular weight range of 32 to 671 has been

19 developed. This development and the compositional evolution of this reference material are described along with an evaluation

20 of the accuracy and stability. This work demonstrates that for the majority of components the accuracy is < 5 % (most < 3 %;

21 < 10 % for hexamethylcyclotrisiloxane (D3-siloxane) and 1,2,4-trichlorobenzene (1,2,4-TCB)) with stabilities of > 2 years (>

22 1 year for acetonitrile, methanol and PFTBA).

23

24 1 Introduction

25 Proton-transfer-reaction mass-spectrometry (PTR-MS) is a technique that allows simultaneous measurements of multiple

26 volatile organic compounds (VOCs) in real-time (≤ 1 s) with high sensitivity (pmol mol⁻¹) and without any sample pre-

27 treatment (Lindinger et al., 1998; Hansel et al., 1999). For these reasons it is a very convenient measurement technique for a

28   wide range of applications. As a result, over the last two decades PTR-MS has developed to become an important and widely

29   applied method for VOC measurements that has resulted in major advances in the field of atmospheric sciences (De Gouw and

30   Warneke, 2007; Park et al., 2013; Yuan et al., 2017). It has also been applied in the medical and food and beverage fields for

31   the detection of VOCs to detect diseases (Beauchamp et al., 2013) and for characterising flavour and odour (Biasioli et al.,

32   2011).

33

34   Multiple manufacturers now produce and commercially sell PTR-MS instruments globally that differ in the production and

35   detection of ions including different types of mass spectrometer. Therefore, as its application becomes more widespread, and

36   more datasets are generated there is an increasing need for accurate calibration and measurement comparability. Additionally,

37   as part of the European funded Aerosols, Clouds and Trace Gases Infrastructure (ACTRIS) project (https://www.actris.eu/)

38   there is an interest to establish PTR-MS as a technique for long-term monitoring of VOCs, which emphasises the need for a

39   robust metrological infrastructure to control and assure the quality of data produced by monitoring stations performing these

40   measurements. However, the lack of traceable reference materials to calibrate PTR-MS instruments presents challenges in the

41   pursuit of obtaining comparable results and is an obstacle for long-term studies. Gas reference materials in high pressure

42   cylinders like those produced by national metrology institutes (NMIs) provide the necessary traceability to the international

43   system of units (SI) ensuring confidence that data is accurate and comparable. These primary reference materials (PRMs)

44   produced by the NMIs represent the highest point in the traceability chain and the accuracy and international comparability is

45   ensured through key comparisons organised within the Consultative Committee on Amount of Substance Gas Analysis

46   Working Group (CCQM-GAWG) and in regional comparisons organised within the Regional Metrology Organisations

47   (RMOs), e.g., EURAMET (Europe).

48

49   As a result of the large number of potential VOCs detectable with PTR-MS it is impractical to have a calibration standard or

50   standards that cover all observable compounds. However, since the conception of PTR-MS, there has been awareness for the

51   potential of this technique to provide quantitative measurements for compounds without the need for specific calibration

52   materials (Hansel et al., 1999). However, in practice, this potential cannot be fully exploited without reliable and applicable

53   methods to retrieve the mass-dependent transmission curve of the instrument. Recent work has demonstrated that a simple

54   reaction kinetics model can accurately predict the sensitivities of different PTR-MS instruments provided the transmission

55   curve was sufficiently constrained (Holzinger et al., 2019). This permits the quantification of measurements of uncalibrated

56   compounds with an accuracy of ≤ 30 %. A prototype PRM to constrain the transmission curve was initially developed for the

57   PTR-MS intercomparison campaign at the CESAR observatory in the central Netherlands (Holzinger et al., 2019), this is NPL

58   PRM A574 reported in this work. Following this comparison exercise improvements to the composition were recommended

59   to include additional compounds in the mass-to-charge (m/Q) 150 – 400 range to provide a more robust retrieval of the mass

60   dependent transmission.

61

62 In this paper, the development and compositional evolution of an NPL PRM specific for constraining the PTR-MS transmission

63 curve is described, including an evaluation of the accuracy through comparisons validating the gravimetric preparation of

64 various different mixtures of similar composition and an assessment of their long-term stabilities.

65 **2 Experimental methods**

66 **2.1 Gravimetric preparation of primary reference materials**

67 The NPL primary reference materials (PRMs) were prepared at four distinct timepoints (September 2017, December-January

68 2018, August 2019 and August 2021) and the compositions evolved over this timeframe (Table 1) due to improvements in the

69 preparation and validation techniques (e.g., 1,2,4-trichlorobenzene; 1,2,4-TCB) or due to requests from the PTRMS

70 community for inclusion of new components (e.g., dimethyl sulphide; DMS). All the NPL PRMs were prepared gravimetrically

71 in accordance with ISO 6142-1:2015 (ISO, 2015) from pure components. All pure components were purity analysed in

72 accordance with ISO 19229 (ISO, 2019). Table 1 provides the sources and purities for each component and shows that all

73 chemicals with the exception of PFTBA were ≥ 98 % pure. All components were liquids at room temperature and pressure,

74 with the exception of propane (gas) and hexamethylcyclotrisiloxane (D3-siloxane; solid). As a solid, D3-siloxane needed to

75 be dissolved into a solvent to enable its addition to the cylinder (see details in the Supporting Information).

76

77 All NPL PRMs were prepared in 10 L aluminium cylinders (Luxfer) with a proprietary passivation treatment (Spectraseal™,

78 BOC) and BS341 no. 15 outlet diaphragm valves (Ceodeux). Cylinders were evacuated using an oil free pump (Scrollvac

79 SC15D, Leybold Vacuum) and turbo molecular pump with magnetic bearing (Turbovac 340M, Leybold Vacuum) to a pressure

80 of $< 3 \times 10^{-7}$ mbar. Individual compounds were added to the evacuated cylinder via a transfer vessel (capped $\frac{1}{8}$" diameter

81 tube, with a nominal volume of 1 mL, Swagelok, electro-polished stainless steel). The transfer vessel was weighed against a

82 tare vessel matched for size and shape before and after each addition into the evacuated cylinder (Mettler-Toledo XP2004S).

83 The ultra-high purity nitrogen balance gas (BIP+, Air Products) was added via direct addition to the cylinder, through purged

84 $\frac{1}{16}$" tubing (Swagelok, electro-polished stainless steel). For the vast majority of compounds, they were initially produced as

85 binary parent mixtures at amount fractions > 10 μmol mol^{-1} (typically at nominally 50 μmol mol^{-1}) though some were produced

86 as ternary or quaternary mixtures containing two or three compounds together in the same parent mixture. A full breakdown

87 of the 50 parent mixtures used for each of the six NPL PRMs are shown in Table 2. Aliquots of each of these parent mixtures

88 were added by direct addition to an evacuated cylinder to produce a final mixture containing all 16 to 20 VOCs at nominal 1

89 μmol mol^{-1} in a balance of nitrogen.

Table 1. Overview of composition (name, formula, CAS# and the monoisotopic molecular ion (MH⁺) formed following protonation in H_3O^+ mode) of the PTRMS transfer curve reference material including the analytical technique used to validate, sources and purities of the pure chemicals and the key comparisons through which the traceability to the international community is derived.

Compound[a]	Formula	CAS #	MH⁺	Traceability (Reference, if applicable)	Validation technique	Source	Purity (%)
methanol	CH_3OH	67-56-1	33.04	EURAMET-1305 OVOCs (Worton et al., 2022), CCQM-K174[c]	Cryotrap GC-FID	Alfa Aesar	>99.9
acetonitrile	CH_3CN	75-05-8	42.05	CCQM-K174[c]	GC-MS	Sigma Aldrich	>99.9
acetaldehyde	CH_3CHO	75-07-0	45.05	CCQM-K174[c]	Cryotrap GC-FID	Fluka	>99.9
propane[b]	C_3H_8	74-98-6	-	CCQM-K111 (Veen et al., 2017)	GC-FID	Air Liquide	>99.9
ethanol	C_2H_5OH	64-17-5	47.07	CCQM-K93 (Brown et al., 2013), EURAMET-1305 OVOCs (Worton et al., 2022), CCQM-K174[c]	GC-FID	Fisher Chemical	>99.9
acetone	CH_3OCH_3	67-64-1	59.08	EURAMET-1305 OVOCs (Worton et al., 2022), CCQM-K174[c]	GC-FID	Alfa Aesar	>99.9
DMS	C_2H_6S	75-18-3	63.13	CCQM-K94 (Lee et al., 2016), CCQM-K165[c]	GC-FID	Alfa Aesar	99.9
isoprene	C_5H_8	78-79-5	69.12	EURAMET-886 (Grenfell et al., 2008; Grenfell et al., 2010)	GC-FID	Aldrich	>99.9
MVK	C_4H_6O	78-94-4	71.09	CCQM-K174[c]	GC-FID	Alfa Aesar	98.0
MEK	C_4H_8O	78-93-3	73.11	CCQM-K174[c]	GC-FID	Sigma Aldrich	>99.9
benzene	C_6H_6	71-43-2	79.11	CCQM-K10.2018 (Cecelski et al., 2022)	GC-FID	Sigma Aldrich	>99.9
toluene	C_7H_8	108-88-3	93.14	CCQM-K10.2018 (Cecelski et al., 2022)	GC-FID	Sigma Aldrich	>99.9
m-xylene	C_8H_{10}	108-38-3	107.16	CCQM-K10.2018 (Cecelski et al., 2022)	GC-FID	Alfa Aesar	99.9
1,2,4-TMB	C_9H_{12}	95-63-6	121.2	EURAMET-886 (Grenfell et al., 2008; Grenfell et al., 2010)	GC-FID	Fluka	99.6
1,2,4-TFB	$C_6H_3F_3$	367-23-7	133.08	-	GC-FID	Alfa Aesar	99.9
3-carene	$C_{10}H_{15}$	13466-78-9	137.24	CCQM-K121 (Liaskos et al., 2018)	GC-FID	Fluka	98.0
1,2,4-TCB	$C_6H_3Cl_3$	120-82-1	182.45	-	GC-FID	Sigma Aldrich	99.5
D3-siloxane	$C_6H_{18}O_3Si_3$	541-05-9	223.46	EURAMET 1305 Siloxanes[c]	GC-FID	Aldrich	99.9
D4-siloxane	$C_8H_{24}O_4Si_4$	556-67-2	297.62	EURAMET 1305 Siloxanes[c]	GC-FID	Aldrich	99.7
D5-siloxane	$C_{10}H_{30}O_5Si_5$	541-02-6	371.77	EURAMET 1305 Siloxanes[c]	GC-FID	Alfa Aesar	99.5
PFTBA	$C_{12}F_{27}N$	311-89-7	672.01	-	GC-FID	Alfa Aesar	87.6

[a]Short names shown here but full IUPAC names are: propan-2-one (acetone), dimethyl sulfide (DMS), 2-Methylbuta-1,3-diene (isoprene), but-3-en-2-one (methyl vinyl ketone; MVK), butan-2-one (methyl ethyl ketone; MEK), 1,2,4-trimethyl benzene (1,2,4-TMB), 1,2,4-trifluoro benzene (1,2,4-TFB), 3,7,7-trimethylbicyclo[4.1.0]hept-3-ene (3-carene), 1,2,4-trichlorobenzene (1,2,4-TCB), hexamethylcyclotrisiloxane (D3-siloxane), octamethylcyclotetrasiloxane (D4-siloxane), decamethylcyclopentasiloxane (D5-siloxane), and perfluoro tertiary butyl amine (PFTBA). [b]Not detectable in PTRMS in H_3O^+ mode but included as an internal standard. [c]Comparison in progress at the time of publication.

Table 2. Composition, hierarchies and parent cylinder IDs (dates prepared) for all components for the six NPL PRMs (A574, A578, A638, A643, 2819 and D961492) prepared in this work.

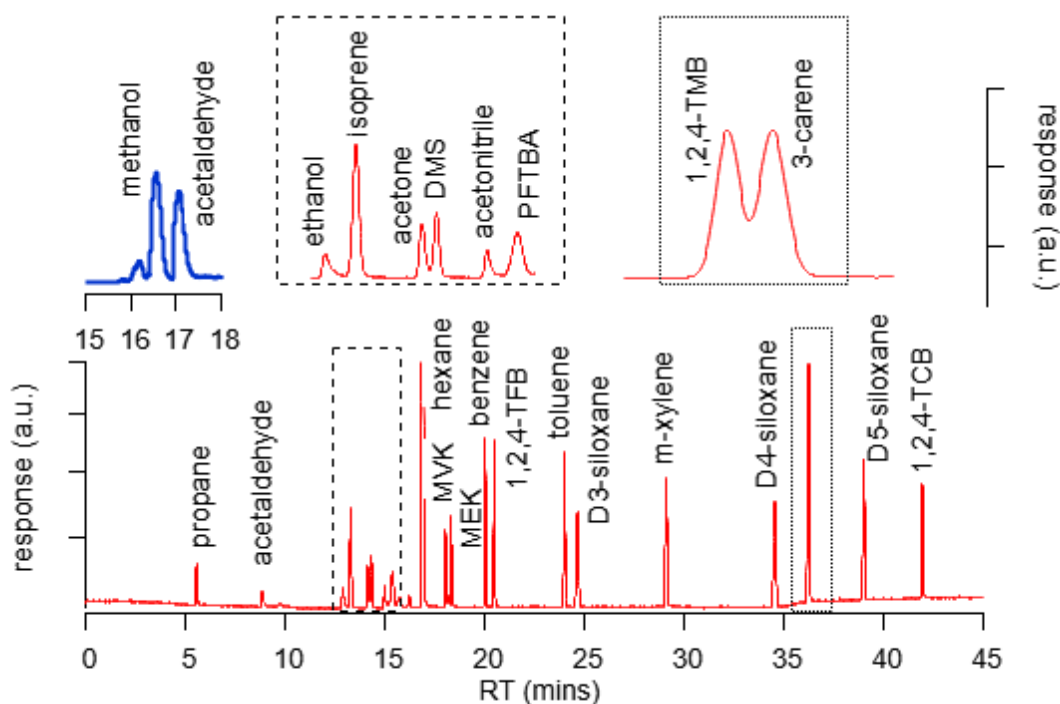
Compound	Cylinder ID (Date Prepared)					
	A574 (18/09/2017)	A578 (18/09/2017)	A638 (04/12/2018)	A643 (02/01/2019)	2819 (23/08/2019)	D961492 (31/08/2021)
methanol		A463 (13/02/2015)			A410 (24/01/2013)	A602 (26/04/2018)
acetonitrile		A389R (11/02/2015)			A403 (11/10/2012)	A670R (01/04/2021)
acetaldehyde		A400R (02/02/2015)			A402 (11/10/2012)	2832 (01/04/2021)
propane		D910381R (18/11/2014)			NG561 (16/10/2014)	-
ethanol		A463 (13/02/2015)			A410 (24/01/2013)	A602 (26/04/2018)
acetone		A463 (13/02/2015)			VOC6 (05/05/2009)	A602 (26/04/2018)
DMS		-	2106 (21/02/2017)		NG388 (13/09/2012)	3073 (16/11/2020)
isoprene		D292194R (13/01/2011)			VOC6 (05/05/2009)	D994138R2 (28/09/2020)
MVK		2064 (24/06/2016)			2088 (24/06/2016)	3070 (02/08/2021)
MEK		A389R (11/02/2015)			A403 (11/10/2012)	3070 (02/08/2021)
benzene		D910381R (18/11/2014)			D842635R (13/10/2015)	D618317 (15/08/2018)
toluene			-			D600070 (19/03/2018)
m-xylene	D641688 (04/03/2010)		D618307 (19/03/2018)		D994171 (20/03/2013)	D618307 (19/03/2018)
1,2,4-TMB	D442684 (02/03/2017)			D711530 (26/11/2018)		D442684 (02/03/2017)
1,2,4-TFB		A569 (11/09/2017)			2810 (21/08/2019)	D723197R (14/06/2021)
3-carene	D090493 (18/11/2014)			D711532 (21/11/2018)		D711532 (21/11/2018)
1,2,4-TCB		-	D641970R (31/07/2018)	A568 (18/08/2017)		D723197R (14/06/2021)
D3-siloxane		-	2586 (07/11/2018)	2693 (07/11/2018)	2586 (07/11/2018)	3134 (22/06/2021)
D4-siloxane	A582 (15/08/2017)	A567 (18/08/2017)			A582 (15/08/2017)	A629R (22/06/2021)
D5-siloxane	A582 (15/08/2017)	A567 (18/08/2017)	A629 (26/11/2018)		A582 (15/08/2017)	A629R (22/06/2021)
PFTBA			-			D961497 (30/07/2021)

2.2 Analytical methods

93   To perform the analytical validation a method was developed on a gas chromatograph (GC; Agilent Technologies 7890)
94   instrument equipped with both a flame ionisation detector (FID) and mass spectrometer (Agilent 5975; GC-FID/MS system).
95   Samples were introduced using a 6 port 2 position valve (VICI) and a fixed sampling loop (1 mL). The column effluent was
96   split to both detectors simultaneously through the use of a detector splitter plate (Agilent Technologies). Separation was
97   achieved for all components using a DB-624 column (J&W; 75m × 0.53 mm, df = 3 µm), although 3-carene and 1,2,4-TMB
98   were not baseline separated. The carrier gas was helium and the flow was held constant at 5 mL min^{-1}, with a temperature
99   program starting at 30 °C held for 10 minutes, ramped at 10 °C min^{-1} to 120 °C and held for 15 minutes before a final ramp of
100   50 °C min^{-1} was applied to a final temperature of 200 °C, which was held for a further 10 minutes. The total run time was 46
101   minutes.

103   Low FID responses for methanol and acetaldehyde presented some challenges for validation work as the observed peaks were
104   too small to achieve useable results due to poor reproducibility. As a result, another analytical method was developed on a
105   second GC-FID instrument without an MS (Scion 456; Cryo-GC-FID system) that had a pre-concentration loop packed with
106   glass beads and cooled with liquid nitrogen that enabled trapping of larger volume samples yielding larger peaks and improved
107   repeatability for these three compounds. Separation was achieved using a Rtx-624 (Restek; 105m × 0.32 mm, df = 1.8 µm).
108   The carrier gas was hydrogen and the flow was held constant at 1 mL min^{-1}, with a temperature program starting at 30 °C held
109   for 5 minutes, ramped at 25 °C min^{-1} to 200 °C with a final hold of 25 minutes. The total run time was 42 minutes.

111   Figure 1 shows the chromatograms obtained from both instruments (cryo-GC-FID, blue; GC-MS/FID, red) and demonstrates
112   that all compounds, with the exception of 1,2,4-trimethyl benzene (1,2,4-TMB) and 3-carene, and acetone and dimethyl sulfide
113   (DMS), were baseline separated. The chromatogram in Fig. 1 shows a valley between the 1,2,4-TMB and 3-carene peaks and
114   between the acetone and DMS peaks that provides sufficient separation to obtain robust and repeatable peak areas for all four
115   compounds.

Figure 1. Chromatogram of NPL PRM 2819 showing separation of compounds in the GC-FID/MS (thin red trace) and cryo-GC-FID (thick blue trace).

2.3 Stability assessment

Stability of all six NPL PRMs were assessed by tracking the ratios of the FID responses of each component relative to an internal reference that was present in every mixture and which is known to be stable (Rhoderick, 2010; Rhoderick and Lin, 2013; Worton et al., 2022). Propane was originally included as an internal tracer to monitor stability but as the PTR-MS in H_3O^+ mode cannot detect this compound it was replaced by benzene. Benzene is a good internal tracer with stability of > 2 years that has been well demonstrated relative to propane and hexane for this cylinder type at 5 µmol mol^{-1} with an uncertainty of 0.5 % (Rhoderick et al., 2019). A similar performance would be expected at 1 µmol mol^{-1} and is demonstrated in this work albeit with an uncertainty of 1 % (Fig. S2, Supporting Information).

All the measurements used for the stability analysis were collected on the same GC-FID/MS instrument with the exception of methanol and acetaldehyde (Cryo-GC-FID). Both instruments remained unchanged throughout the entire time-period of the measurements, which spanned more than 4 years. The observed responses for each compound were corrected for differences in the gravimetric amount fraction and ratioed against the internal reference compound benzene, that was present in every mixture. The uncertainties in the observed ratios included uncertainties for the gravimetric preparation and the repeatability of

135   the analyses. The combined standard uncertainties were multiplied by a coverage factor of 2 ($k=2$) providing a coverage

136   probability of 95 %. The observed ratios were normalised to the average response of all data for that compound to enable

137   comparisons between compounds with different FID responses. For this analysis all the data for all six NPL PRMs were

138   considered together to enable an understanding of stability across a longer time period than would be possible for a single

139   PRM. Least squares fit straight-line regressions were modelled to the temporal evolution of the data to determine if there was

140   any statistically significant change in amount fraction of any of the compounds in the NPL PRMs. The slopes from these

141   regression analyses were evaluated with an analysis of variance (ANOVA) test using the 'StatsLinearRegression' function in

142   IGOR pro 8.04 (Wavemetrics) (Zar, 1999; Snedecor and Cochran, 1989) to determine whether they were significantly different

143   to zero using an *F*-test, i.e., no drift in amount fraction during the measurement period (F < Fc).

144

2.4 Validation approach

146   All 5 NPL PRMs (A574, A578, A638, A643, D961492) were validated against NPL PRM 2819 with the exception of PFTBA

147   and toluene that were only present in the most recent NPL PRM (D961492). NPL PRM 2819 was used as the reference for all

148   the validations because the parents used for the preparation of this mixture were deliberately different from all other mixtures

149   with the explicit goal of enabling the most robust validation. All compounds were analysed on the GC-FID/MS system, with

150   the exception of methanol and acetaldehyde (Cryo-GC-FID). Toluene was validated by comparison against an existing NPL

151   PRM containing BTEX (benzene, toluene, *m*-xylene, *p*-xylene and *o*-xylene) components. PFTBA was validated against the

152   gravimetric data used to make two independent certified reference materials. The majority of the validation work took place

153   between September and December 2020 with one in 2019 and 2022, respectively, and three in 2021 (see Table S1, Supporting

154   Information). As such there is an influence of stability on the validation data as the PTRMS NPL PRMs differed in age at the

155   time of validation.

156

157   Each comparison was conducted by running the NPL PRMs (A574, A578, A638, A643 and D961492) against NPL PRM 2819

158   in a repeating alternating pattern, (ABA)$_n$ where A represents NPL PRM 2819 and B one of the other PRMs (*j*) and with the

159   number of repeats ranging between 3 and 5 (n = 3 − 5). The ratio in response was determined by bracketing between the nearest

160   neighbours and the average ratio was calculated for each compound based on the number of repeats along with the associated

161   standard deviation. The assigned analytical value for compound *i* in NPL PRM *j* ($x_{u,i,j}$) was calculated by multiplying the

162   average ratio by the gravimetrical amount fraction ($x_{s,i}$) of compound *i* in NPL PRM 2819. The relative difference (Δx) between

163   the assigned analytical value and the gravimetric value of compound *i* in NPL PRM *j* was calculated from:

164

165   $$\Delta x = \frac{(x_{u,i,j} - x_{s,i})}{x_{s,i}} \times 100 \qquad\qquad\qquad\qquad \text{Eq. 1}$$

166

167 The uncertainty in the relative difference combined the standard uncertainty in the repeatability in the analysis with the
168 gravimetric uncertainty. The combined standard uncertainty was multiplied by a coverage factor of 2 (k=2) providing a
169 coverage probability of 95 %.

170 **3 Results**

171 **3.1 Composition**

172 The PTR-MS transfer curve reference material contains 20 different VOCs that span a wide range of molecular masses from
173 32 to 671 and a range of functional group classes including alcohol, aldehyde, ketone, alkene, aromatic, halocarbon and
174 siloxane (Table 2). With PTR-MS, most are entirely detected at their protonated mass, as well as a few compounds that partially
175 fragment during protonation (e.g. monoterpenes, siloxanes, and isoprene). The compounds were chosen by considering the
176 needs of the PTR-MS user community to cover the full range of mass-to-charge ratios (m/Q) encountered, their low
177 fragmentation following proton transfer and because many are of relevance in atmospheric measurements, which was the initial
178 intended target end user group. Other compounds were included as a consequence of the preparation method, that is the case
179 of *n*-hexane included as diluent for D3-siloxane or propane which was present in one of the premixtures and was originally
180 included as an internal tracer to monitor stability. The composition evolved over time, as shown in Table 2, with DMS, 1,2,4-
181 TCB, D3-siloxane, toluene and PFTBA being added at different times, and propane being removed in the final iteration. For
182 D4-siloxane there was a preparation error, and it was not added to either NPL PRM A638 or A643.

183

184 An amount fraction of nominally 1 μmol mol^{-1} in a balance gas of nitrogen was selected as a compromise between preparation
185 complexity and mixture stability. This amount fraction enabled many components to be prepared from parent mixtures of
186 higher amount fraction (≥10 μmol mol^{-1}), which substantially simplifies the preparation process. This amount fraction also
187 provided a reasonable starting point for stability of the wide range of function groups present in the mixture some of which are
188 known to have more limited stability at lower abundances fractions (nmol mol^{-1}) (Allen et al., 2018).

189 **3.2 Traceability to the International System of Units (SI)**

190 Traceability of the primary realisations to the international community through CCQM key comparisons or regional
191 EURAMET comparisons provides confidence in the accuracy of the amount fractions for all components. Table 1 shows which
192 comparisons underpin the traceability for each of the different components. All the components are underpinned by at least
193 one CCQM or EURAMET comparison with the exception of 1,2,4-TFB, 1,2,4-TCB and PFTBA, for which there are currently
194 no existing relevant comparisons.

3.3 Hierarchies

Table 2 shows all the parent mixtures and their preparation dates used for the preparation of all 6 NPL PRMs. In total, 50 different parent mixtures were used to produce the 6 NPL PRMs (A574, A578, A638, A643, 2819, D961492). In general, parent mixtures were similar for A574, A578, A638 and A638 but were different to 2819 and D961492 providing independence and thus confidence in the validation work and thus the preparations. There were a few exceptions. For m-xylene the parent used for A638 and A643 were the same as D961492 but different to A574, A578 and D961492. For 1,2,4-TMB only two parent mixtures were used one for A574, A578 and D961492 and another for A638, A643 and 2819. For 3-carene only two parents were used one for A574 and A578 and another for A638, A643, 2819 and D961492. For D3-siloxane three parents were used, one for A638 and 2819, another for A643 and another for D961492.

3.4 Validation

Figure 2 shows the relative differences (Δx) determined from Eq. 1 for all compounds using all the validation data obtained from the 13 comparisons outlined in Table S1 (Supporting Information). All the data shown in Fig. 2 is the FID data from the GC-MS/FID instrument with the exception of acetonitrile (MS data from the GC-MS/FID instrument), methanol and acetaldehyde (FID data from the cryo-GC-FID instrument). The MS data is used for acetonitrile because the FID data shows more variability likely as a result of the small FID peak size. This variability is not observed in the MS data providing better precision (Fig. S1, Supporting Information).

In general, the data from Fig. 2 could be split into three groups. The first group consisted of propane, isoprene, benzene, toluene, 3-carene, methanol, acetonitrile, acetaldehyde, *m*-xylene, 1,2,4-TMB, MEK where the spread in the validation data is within 3 % and these represent components where NPL had substantial prior experience. The second group is acetone, DMS, MVK and PFTBA where the spread in the validation data is within 5 % and these are relatively new components where capabilities were developed more recently. Recognising the challenges in preparing PRMs containing siloxanes as a result of their low vapour pressures and observing the recent improvements in preparation since 2019, the D4-siloxane and D5-siloxanes can also be categorised as group 2 after excluding the earliest parent preparations used for A574 and A578 in 2017, which are inconsistent with more recent work as part of the EURAMET 1305 Siloxanes comparison (Van Der Veen et al., 2022). The final group is comprised of D3-siloxane and 1,2,4-TCB where the spread in validation data is within 10 % and these compounds represent those which the most challenging to prepare as a result of their low vapour pressures. There is an observable bias of about 8 % between two groups of mixtures; one group is A638 and 2819 and the other is A643, D961492. This reflects differences between the parent mixtures (2586, 2693 and 3134) that resulted from the challenges in preparation. Ethanol also sits with this group in part due to the small size of the peak observed in the GC-MS/FID instrument and because of what looks like an outlier (A643), suggesting some potential losses during preparation that were unique to this one NPL PRM.

226 **Figure 2.** Relative difference (Δx) using the FID data (except acetonitrile, which uses the MS data) for all components in the

227 different primary reference materials (A574, A578, A638, A643 and D961492) relative to primary reference material 2819

228 (solid symbols). The solid black line represents the average of these validations with the error bar representing the associated

229 expanded uncertainty (2σ). For D4- and D5-siloxane the averages do not include the validations from A574 or A578. Methanol

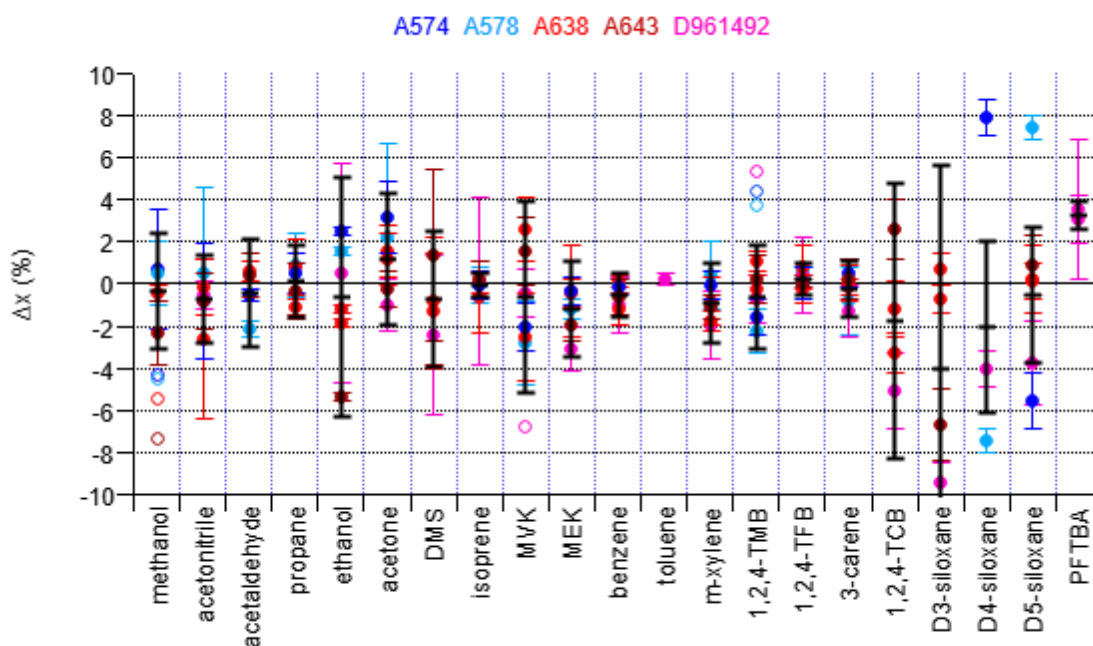
230 and acetaldehyde data are from the cryo-GC-FID instrument while all others are from the GC-FID/MS instrument. Open

231 symbols represent the original data before correcting for biases observed in three of the parent mixtures (A410, 5 % low for

232 methanol; 3070, 6.3 % low for MVK and D711530 6 % low for 1,2,4-TMB). PFTBA and toluene were only included in the

233 most recently prepared PRM (D961492) and are not present in 2819. Their validation is described in the text. Supporting

234 validation data from all the MS and FID measurements is shown in Fig. S1 (Supporting Information).

235

236

237

238 All the FID and supporting MS data for all compounds are shown in Fig. S1 (Supporting Information). No MS data was

239 available for toluene, 1,2,4-TCB or PFTBA because the relevant single m/Q ions had not been included in the MS single ion

240 monitoring method at the time of analysis and methanol where the MS signal was too small to provide a reliable response.

241 Figure S1 shows very good agreement between the FID and MS validation with all components agreeing within the

242 uncertainties providing confidence in the validation results.

243

244 In addition to the observed bias in parent mixtures for D3-siloxane three other parent mixtures were also discovered to be

245 biased after re-analysis. The observed differences have been corrected for in Fig. 2 and Fig. S1 (Supporting Information). For

246 methanol, one parent (A410) was confirmed to be 5.0 % high relative to the other parents (A463, A540 and A602) For MVK,

247 one parent (3070) was confirmed to be 6.3 % low relative to the other parents (2064 and 2088). For 1,2,4-TMB, one parent

248 D711530 was confirmed to be 6.0 % low relative to D442684 and other in-house standards of 1,2,4-TMB not used in this work

249 but used to prepare 30 component ozone precursor mixtures at NPL (Grenfell et al., 2010).

250 **3.5 Stability**

251 Figure 3 shows the stability data for four selected compounds; methanol, isoprene, D3-siloxane and PFTBA. These were

252 selected as representative examples of the different observed stability behaviours although the stability data plots

253 corresponding to all compounds are shown in Fig. S2 (Supporting Information). The trendlines from the least squares fit

254 straight-line regressions shown in Fig. 3 and Fig. S2 (Supporting Information) were used to determine the annual drift rates

255 shown in Table 3 and Fig. 4. From the ANOVA test there are statistically significant trends ($F > F_c$) for 10 of the compounds

256 (methanol, acetonitrile, acetaldehyde, ethanol, acetone, DMS, isoprene, MVK, benzene, D3-siloxane) but these trends are

257 small (< 2 % yr^{-1}) except for methanol, acetonitrile and PFTBA.

258

259 Methanol and acetaldehyde were the only two components that were measured on the cryo-GC-FID and hence the datasets are

260 more limited. A result is that there is no overlap between the three NPL PRMs so any systematic differences between them

261 may result in an artificial bias, which may exacerbate any stability trend. More work is needed to confirm this. The stability

262 data for D3-siloxane reflects the observed validation bias and shows two clear trends; one for A638 and 2819 and the other for

263 A643 and D961492. For the regression analysis and drift calculations these have been treated independently (Table 3).

264

265 All compounds, with the exceptions of methanol, acetonitrile and PFTBA, show trends similar to isoprene with good stability

266 and annual drift rates of < 3 % yr^{-1} (Table 3). For acetonitrile the large spread in validation data (FID data; Fig. S1, Supporting

267 Information) leads to a noisy stability dataset that may play a role in the larger observed drift rate or this component maybe

268 less stable. Additional work is needed to confirm. As PFTBA was only included in the last NPL PRM (D961492) the stability

269 data only represents about half a year and extrapolating the current trend to 1 and 2 years results in a drift rate that is not

270 accurate as interpolation of the data would suggest no statistical change in amount fraction and minimal drift. More data is

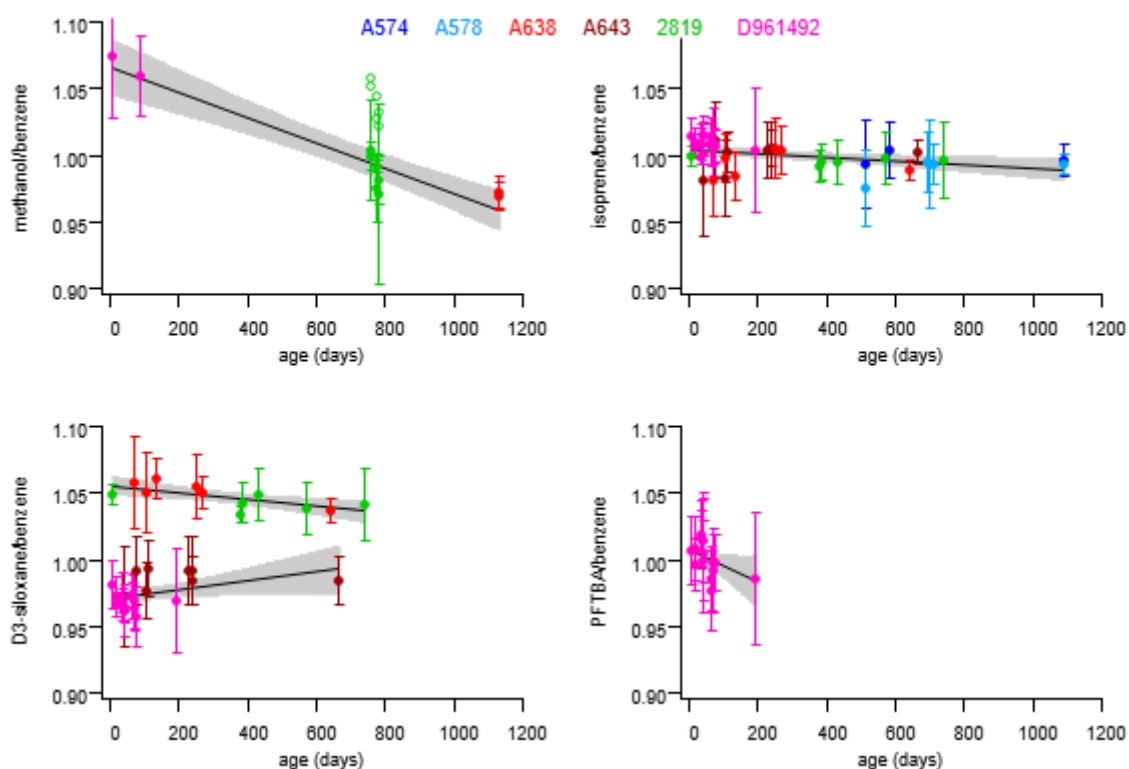
271 needed to confirm the longer-term stability behaviour of PFTBA.

272

273

274

275

276 **Figure 3.** Stability of normalised response with time for four selected compounds relative to benzene, methanol (top left),
277 isoprene (top right), D3-siloxane (bottom left) and PFTBA (bottom right) for all six NPL PRMs (solid symbols). The open
278 symbols (methanol; top left) show the original data before being corrected for an observed 5.0 % bias in the parent mixture
279 (A410). The best fit curves from least squares straight line regression analyses are shown (solid black line) along with the 95
280 % confidence interval of the fits (shaded area). The slope, intercept and F-statistic data from the regression analyses are shown
281 in Table 3. Stability plots for all compounds are shown in Fig. S2 (Supporting Information).
282

283
284
285
286 Given the age differences between the different PRMs at the time of validation (233 – 709 days; Table S1, Supporting
287 Information) it is not possible to deconvolute the contributions of stability and preparation to the observed validation
288 differences. However, Fig. 4 shows that for the majority of compounds there is good agreement between the observed average
289 validation data and the calculated drift for over 1 – 2 years, with the exception of methanol, acetonitrile and PFTBA, which
290 differ for the reasons discussed previously. These observations are consistent with the age differences of the different PRMs
291 at the time of validation indicating that stability was likely the major driver between the observed validation differences.
292

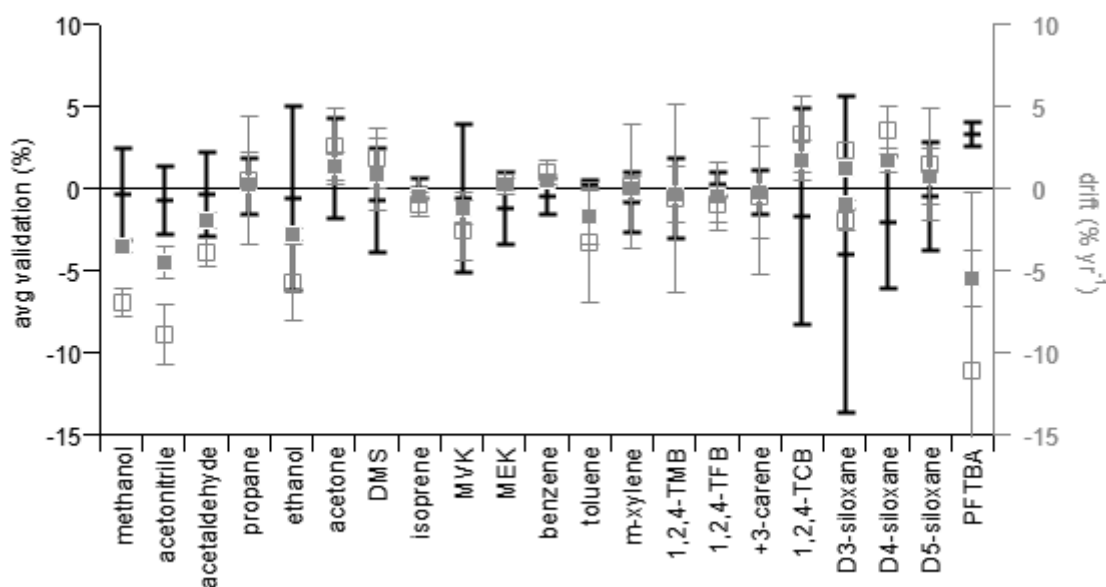
293 **Table 3.** Summary of the results of the least squares straight-line regression analysis for all stability data shown in Figures 3
294 and S2 (Supporting Information). Results are shown for the slope (± 2σ), intercept (± 2σ), ANOVA test statistics (F and Fc)
295 used to evaluate the presence of a statistically significant trend (F > Fc), the calculated annual drift (± 2σ) determined from the
296 linear fit and the average of the validation data (± 2σ), also shown in Figure 2.

297

compound	slope (× 10^{-5})	intercept	F	Fc	Drift (%/yr)	Avg valid. (%)
methanol[1]	-9.539 ± 2.700	1.067 ± 0.021	57.005	5.318	-3.48 ± 0.42	-0.37 ± 2.77
acetonitrile	-12.328 ± 6.128	1.036 ± 0.026	16.530	4.085	-4.50 ± 0.94	-0.72 ± 2.08
acetaldehyde[1]	-5.345 ± 2.800	1.037 ± 0.022	19.699	5.318	-1.95 ± 0.44	-0.40 ± 2.53
propane	0.653 ± 5.393	0.997 ± 0.027	0.062	4.225	0.24 ± 1.97	0.16 ± 1.66
ethanol	-7.841 ± 7.55	1.023 ± 0.032	4.405	4.085	-2.86 ± 0.36	-0.61 ± 5.64
acetone	3.462 ± 3.206	0.990 ± 0.013	4.765	4.085	1.26 ± 0.86	1.18 ± 3.08
DMS	2.441 ± 2.351	0.995 ± 0.007	4.473	4.149	0.89 ± 2.24	-0.76 ± 3.22
isoprene	-1.338 ± 0.975	1.004 ± 0.004	7.690	4.085	-0.49 ± 1.17	-0.04 ± 0.60
MVK	-3.523 ± 2.564	1.010 ± 0.011	7.708	4.085	-1.29 ± 0.94	-0.61 ± 4.50
MEK	0.575 ± 1.967	0.998 ± 0.008	0.349	4.085	0.21 ± 0.36	-1.23 ± 2.25
benzene[2]	1.329 ± 0.983	0.996 ± 0.004	7.456	4.085	0.49 ± 0.18	-0.48 ± 1.07
toluene[3]	-3.546 ± 4.536	1.002 ± 0.004	2.902	4.747	-1.30 ± 1.66	0.19 ± 0.29
m-xylene	0.129 ± 2.034	1.000 ± 0.009	0.016	4.085	0.05 ± 0.74	-0.87 ± 1.88
1,2,4-TMB	-0.870 ± 5.155	1.003 ± 0.022	0.116	4.085	-0.32 ± 1.69	-0.57 ± 2.42
1,2,4-TFB	-1.373 ± 1.448	1.004 ± 0.006	3.672	4.085	-0.50 ± 2.05	0.27 ± 0.74
+3-carene	-0.734 ± 4.631	1.002 ± 0.019	0.103	4.085	-0.27 ± 2.84	-0.25 ± 1.33
1,2,4-TCB	4.512 ± 6.455	0.991 ± 0.018	2.027	4.149	1.65 ± 1.16	-1.73 ± 6.56
D3-siloxane[3]	-2.641 ± 1.740	1.056 ± 0.007	11.444	4.965	-0.96 ± 0.29	-4.02 ± 9.67
	3.195 ± 3.220	0.970 ± 0.007	4.287	4.351	1.17 ± 0.56	
D4-siloxane[4]	4.799 ± 4.300	0.988 ± 0.012	0.765	4.225	1.75 ± 0.74	-2.03 ± 4.06
D5-siloxane[4]	2.066 ± 0.390	0.985 ± 0.026	1.833	4.085	0.75 ± 1.68	-0.49 ± 3.27
PFTBA[5]	-12.045 ± 13.440	1.007 ± 0.010	3.813	4.747	-4.40 ± 1.66	3.31 ± 0.70

298 [a]The GC-FID data for methanol and acetaldehyde was too small to be quantified so this data is from the cryo-GC-FID data and is limited. [b]Benzene stability
299 was determined relative to isoprene. [c]There was a clear bias between several of the PRMs caused by differences in the parent mixtures used so the trends were
300 fitted to the two obvious groupings. [d]Data from A574 and A578 was excluded from the regression analysis. [e]Toluene and PFTBA were only included in the
301 most recent NPL PRM so the assessment of stability is limited in its duration to only 200 days.

302

Figure 4. Comparison of 1-year (filled grey squares) and 2-year (open grey squares) drift rates, calculated from the data in Table 3, with the average validation data (black bars) taken from Fig. 2. For D3-siloxane there are two datapoints for the drift correspond to the two regressions shown in Table 3. The error bars represent the associated expanded uncertainties, representing the 95 % confidence limit.

4 Conclusions

In this work the development of a new primary reference material (PRMs) for constraining the mass dependent transmission curve of PTR-MS instruments has been described along with an evaluation of their validation and stability. Six of these PRMs have been prepared to date from a suite of 50 parent mixtures. In general, there is evidence of very good agreement for the majority of components that supports the robustness of the preparation and 2 years of stability. Challenges were observed in preparation for the least volatile compounds especially for D3-siloxane due to it being a solid at room temperature and pressure. More work is needed to better describe the long-term stability of methanol, acetonitrile and PFTBA. This work highlighted several challenges in analysis that could be resolved by the development of a new analytical method utilised a single instrument equipped with both a preconcentration trap and dual detector setup (MS and FID).

319 Future work to improve the uncertainty of individual components that have the greatest influence on the transmission curve fit

320 would have the biggest influence on the accuracy and repeatability of the transfer curve fit thus maximising the impact of

321 future improvements for the PTRMS user community. For PTRMS instruments that utilise time of flight mass spectrometers

322 the focus would be on improving the uncertainty of the largest molecular weight components specifically the D3-, D4-, D5-

323 siloxanes and 1,2,4-TCB, which represent the greatest challenges in preparation due to their low vapour pressures.

324 **Data availability**

325 All data used to produce the figures in this paper are available on request.

326 **Competing interests**

327 The authors declare that they have no conflict of interest.

328 **Author contributions**

329 DRW and RH conceptualized the work. DRW processed the data, produced the figures and tables and wrote the paper. SM

330 developed novel methods for the preparation of primary and certified reference materials, planned and prepared all reference

331 materials, conducted all the validation and stability data collection and contributed to the writing of the paper. KOD contributed

332 to the preparation of reference materials, reprocessed some of the stability data and worked with SM to prepare the certified

333 reference materials. RH provided inputs to define the composition of the PTR-MS reference material and reviewed the paper.

334 **Financial support**

335 NPL acknowledges funding from the Department of Business Energy and Industrial Strategy (BEIS) National Measurement

336 System. RH acknowledges funding from the EMPIR programme (19ENV06 MetClimVOC), co-financed by the Participating

337 States and from the European Union's Horizon 2020 research and innovation programme, and from the European Union's

338 Horizon 2020 research and innovation programme (ACTRIS-2) under grant agreement no. 654109.

339 **References**

340 Allen, N. D. C., Worton, D. R., Brewer, P. J., Pascale, C., and Niederhauser, B.: The importance of cylinder passivation for
341   preparation and long-term stability of multicomponent monoterpene primary reference materials, Atmos. Meas. Tech., 11,
342   6429-6438, 10.5194/amt-11-6429-2018, 2018.
343 Beauchamp, J., Herbig, J., Dunkl, J., Singer, W., and Hansel, A.: On the performance of proton-transfer-reaction mass
344   spectrometry for breath-relevant gas matrices, Measurement Science and Technology, 24, 125003, 2013.

345 Biasioli, F., Gasperi, F., Yeretzian, C., and Märk, T. D.: PTR-MS monitoring of VOCs and BVOCs in food science and
346   technology, TrAC Trends in Analytical Chemistry, 30, 968-977, https://doi.org/10.1016/j.trac.2011.03.009, 2011.
347 de Gouw, J. and Warneke, C.: Measurements of volatile organic compounds in the earths atmosphere using proton-transfer-
348   reaction mass spectrometry, Mass Spectrom. Rev., 26, 223-257, 10.1002/mas.20119, 2007.
349 Grenfell, R. J. P., Milton, M. J. T., Harling, A. M., Vargha, G. M., Brookes, C., Quincey, P. G., and Woods, P. T.: Standard
350   mixtures of ambient volatile organic compounds in synthetic and whole air with stable reference values, Journal of
351   Geophysical Research-Atmospheres, 115, 10.1029/2009jd012933, 2010.
352 Hansel, A., Jordan, A., Warneke, C., Holzinger, R., Wisthaler, A., and Lindinger, W.: Proton-transfer-reaction mass
353   spectrometry (PTR-MS): on-line monitoring of volatile organic compounds at volume mixing ratios of a few pptv, Plasma
354   Sources Science & Technology, 8, 332-336, 10.1088/0963-0252/8/2/314, 1999.
355 Holzinger, R., Acton, W. J. F., Bloss, W. J., Breitenlechner, M., Crilley, L. R., Dusanter, S., Gonin, M., Gros, V., Keutsch, F.
356   N., Kiendler-Scharr, A., Kramer, L. J., Krechmer, J. E., Languille, B., Locoge, N., Lopez-Hilfiker, F., Materić, D., Moreno,
357   S., Nemitz, E., Quéléver, L. L. J., Sarda Esteve, R., Sauvage, S., Schallhart, S., Sommariva, R., Tillmann, R., Wedel, S.,
358   Worton, D. R., Xu, K., and Zaytsev, A.: Validity and limitations of simple reaction kinetics to calculate concentrations of
359   organic compounds from ion counts in PTR-MS, Atmos. Meas. Tech., 12, 6193-6208, 10.5194/amt-12-6193-2019, 2019.
360 ISO: 6142-1:2015 Gas analysis - Preparation of calibration gas mixtures - Gravimetric method for class I mixtures.
361   International Organization for Standardisation (ISO), Geneva, Switzerland, 2015.
362 ISO: 19229:2019 Gas analysis - Purity analysis and the treatment of purity data. International Organization for Standardisation
363   (ISO), Geneva, Switzerland, 2019.
364 Lindinger, W., Hansel, A., and Jordan, A.: On-line monitoring of volatile organic compounds at pptv levels by means of
365   proton-transfer-reaction mass spectrometry (PTR-MS) medical applications, food control and environmental research,
366   International Journal of Mass Spectrometry and Ion Processes, 173, 191-241, http://dx.doi.org/10.1016/S0168-
367   1176(97)00281-4, 1998.
368 Park, J. H., Goldstein, A. H., Timkovsky, J., Fares, S., Weber, R., Karlik, J., and Holzinger, R.: Eddy covariance emission and
369   deposition flux measurements using proton transfer reaction – time of flight – mass spectrometry (PTR-
370   TOF-MS): comparison with PTR-MS measured vertical gradients and fluxes, Atmos. Chem. Phys., 13, 1439-1456,
371   10.5194/acp-13-1439-2013, 2013.
372 Rhoderick, G. C.: Stability assessment of gas mixtures containing terpenes at nominal 5 nmol/mol contained in treated
373   aluminum gas cylinders, Analytical and Bioanalytical Chemistry, 398, 1417-1425, 10.1007/s00216-010-4058-0, 2010.
374 Rhoderick, G. C. and Lin, J.: Stability Assessment of Gas Mixtures Containing Monoterpenes in Varying Cylinder Materials
375   and Treatments, Analytical Chemistry, 85, 4675-4685, 10.1021/ac400324v, 2013.
376 Rhoderick, G. C., Cecelski, C. E., Miller, W. R., Worton, D. R., Moreno, S., Brewer, P. J., Viallon, J., Idrees, F., Moussay, P.,
377   Kim, Y. D., Kim, D., Lee, S., Baldan, A., and Li, J.: Stability of gaseous volatile organic compounds contained in gas
378   cylinders with different internal wall treatments, Elementa-Sci. Anthrop., 7, 22, 10.1525/elementa.366, 2019.
379 Snedecor, G. W. and Cochran, W. G.: Statistical methods, Iowa State University Press1989.
380 Van der Veen, A. M. H., de Krom, I., Nieuwenkamp, G., Culleton, L. P., Worton, D. R., and Allden, J.: Bilateral comparison
381   VSL-NPL on primary gas standards for the siloxane content in biomethane, TBD, 2022.
382 Worton, D. R., Moreno, S., Brewer, P. J., Li, J., Baldan, A., and van der Veen, A. M. H.: Bilateral comparison of primary
383   reference materials (PRMs) containing methanol, ethanol and acetone in nitrogen, Accreditation and Quality Assurance,
384   10.1007/s00769-022-01513-y, 2022.
385 Yuan, B., Koss, A. R., Warneke, C., Coggon, M., Sekimoto, K., and de Gouw, J. A.: Proton-Transfer-Reaction Mass
386   Spectrometry: Applications in Atmospheric Sciences, Chemical Reviews, 117, 13187-13229,
387   10.1021/acs.chemrev.7b00325, 2017.
388 Zar, J. H.: Biostatistical Analysis, 4th, Prentice Hall, Englewood Cliffs, New Jersey1999.