HAL
open science

ACTRIS non-methane hydrocarbon intercomparison experiment in Europe to support WMO GAW and EMEP observation networks

C. Hoerger, A. Claude, C. Plass-Duelmer, S. Reimann, E. Eckart, R. Steinbrecher, J. Aalto, J. Arduini, N. Bonnaire, N. Cape, et al.

To cite this version:

C. Hoerger, A. Claude, C. Plass-Duelmer, S. Reimann, E. Eckart, et al.. ACTRIS non-methane hydrocarbon intercomparison experiment in Europe to support WMO GAW and EMEP observation networks. Atmospheric Measurement Techniques, 2015, 8 (7), pp.2715 - 2736. 10.5194/amt-8-2715-2015 . hal-01806123

HAL Id: hal-01806123

https://hal.science/hal-01806123v1

Submitted on 27 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Atmos. Meas. Tech., 8, 2715–2736, 2015
www.atmos-meas-tech.net/8/2715/2015/
doi:10.5194/amt-8-2715-2015
© Author(s) 2015. CC Attribution 3.0 License.

Atmospheric Measurement Techniques
Open Access

ACTRIS non-methane hydrocarbon intercomparison experiment in Europe to support WMO GAW and EMEP observation networks

C. C. Hoerger[1], A. Claude[2], C. Plass-Duelmer[2], S. Reimann[1], E. Eckart[3], R. Steinbrecher[3], J. Aalto[4], J. Arduini[5], N. Bonnaire[6], J. N. Cape[7], A. Colomb[8], R. Connolly[9], J. Diskova[10], P. Dumitrean[11], C. Ehlers[12], V. Gros[6], H. Hakola[13], M. Hill[1], J. R. Hopkins[14], J. Jäger[12], R. Junek[15], M. K. Kajos[16], D. Klemp[12], M. Leuchner[17], A. C. Lewis[14], N. Locoge[18], M. Maione[5], D. Martin[19,20], K. Michl[2], E. Nemitz[7], S. O'Doherty[17], P. Pérez Ballesta[9], T. M. Ruuskanen[16], S. Sauvage[18], N. Schmidbauer[21], T. G. Spain[20], E. Straube[15], M. Vana[10,22], M. K. Vollmer[1], R. Wegener[12], and A. Wenger[1,19]

[1]Swiss Federal Laboratories for Materials Science and Technology (Empa), Dübendorf, Switzerland
[2]Deutscher Wetterdienst, Hohenpeissenberg, Germany
[3]Karlsruhe Institute of Technology (KIT), Campus Alpin, Garmisch-Partenkirchen, Germany
[4]Department of Forest Sciences, University of Helsinki, Helsinki, Finland
[5]Department of Environmental Science, University of Urbino "Carlo Bo", Urbino, Italy
[6]Laboratoire des Sciences du Climat et de l'Environnement, Gif sur Yvette, France
[7]Centre for Ecology & Hydrology, Penicuik, UK
[8]Observatoire de Physique du Globe de Clermont-Ferrand, Aubière, France
[9]European Commission Joint Research Centre, Ispra, Italy
[10]Czech Hydrometeorological Institute, Prague, Czech Republic
[11]Ricardo-AEA, Oxon, UK
[12]Institute of Energy and Climate Research, Troposphere (IEK-8), Forschungszentrum Jülich, Jülich, Germany
[13]Finnish Meteorological Institute, Helsinki, Finland
[14]National Centre for Atmospheric Science (NCAS), University of York, York, UK
[15]Umweltbundesamt, Station Schmücke, Gehlberg, Germany
[16]Department of Physics, University of Helsinki, Helsinki, Finland
[17]Fachgebiet für Ökoklimatologie, Technische Universität München, Freising, Germany
[18]Ecole des Mines de Douai, Douai, France
[19]School of Chemistry, University of Bristol, Bristol, UK
[20]School of Physics and Centre for Climate and Air Pollution Studies, National University of Ireland, Galway, Ireland
[21]Norwegian Institute for Air Research (NILU), Kjeller, Norway
[22]Global Change Research Centre, Academy of Sciences of the Czech Republic, Brno, Czech Republic

Correspondence to: S. Reimann (stefan.reimann@empa.ch)

Received: 17 September 2014 – Published in Atmos. Meas. Tech. Discuss.: 13 October 2014
Revised: 06 May 2015 – Accepted: 23 May 2015 – Published: 08 July 2015

Abstract. The performance of 18 European institutions involved in long-term non-methane hydrocarbon (NMHC) measurements in ambient air within the framework of the Global Atmosphere Watch (GAW) and the European Monitoring and Evaluation Programme (EMEP) was assessed with respect to data quality objectives (DQOs) of ACTRIS (Aerosols, Clouds, and Trace gases Research InfraStructure Network) and GAW. Compared to previous intercomparison studies the DQOs define a novel approach to assess and ensure a high quality of the measurements. Having already been adopted by GAW, the ACTRIS DQOs are demanding with deviations to a reference value of less than 5 % and a repeatability of better than 2 % for NMHC mole fractions above 0.1 nmol mol^{-1}.

Published by Copernicus Publications on behalf of the European Geosciences Union.

The participants of the intercomparison analysed two dry gas mixtures in pressurised cylinders, a 30-component NMHC mixture in nitrogen (NMHC_N_2) at approximately 1 nmol mol^{-1} and a whole air sample (NMHC_air), following a standardised operation procedure including zero- and calibration gas measurements. Furthermore, participants had to report details on their instruments and assess their measurement uncertainties.

The NMHCs were analysed either by gas chromatography–flame ionisation detection (GC-FID) or by gas chromatography–mass spectrometry (GC-MS). For the NMHC_N_2 measurements, 62 % of the reported values were within the 5 % deviation class corresponding to the ACTRIS DQOs. For NMHC_air, generally more frequent and larger deviations to the assigned values were observed, with 50 % of the reported values within the 5 % deviation class. Important contributors to the poorer performance in NMHC_air compared to NMHC_N_2 were a more complex matrix and a larger span of NMHC mole fractions (0.03–2.5 nmol mol^{-1}). The performance of the participating laboratories were affected by the different measurement procedures such as the usage of a two-step vs. a one-step calibration, breakthroughs of C_2–C_3 hydrocarbons in the focussing trap, blank values in zero-gas measurements (especially for those systems using a Nafion® Dryer), adsorptive losses of aromatic compounds, and insufficient chromatographic separation.

1 Introduction

Volatile organic compounds (VOCs) are important atmospheric trace gases with anthropogenic and biogenic emissions (e.g. Koppmann, 2007; Warneck, 1988, and references therein). VOCs include a large variety of non-methane hydrocarbons (NMHCs, mostly from C_2–C_{16}) such as alkanes, alkenes, alkynes, aromatic compounds, and terpenoids as well as oxygenated VOCs (OVOCs) such as alcohols, aldehydes, and ketones (Andreae and Merlet, 2001; Monks et al., 2009; Placet et al., 2000; Plass-Duelmer et al., 1993; Sawyer et al., 2000). The mole fractions of these compounds vary from below 1 pmol mol^{-1} to tens of nmol mol^{-1} in background and urban air, respectively (e.g. Gros et al., 2007; Parrish and Fehsenfeld, 2000). Atmospheric VOCs have an impact on the oxidising capacity of the atmosphere through their role in the generation of photo-oxidants (e.g. ozone and organic radicals) and are precursors of secondary organic aerosols. For these reasons, reliable measurements of VOCs are essential, and they are consequently included in the long-term monitoring programmes of the Global Atmosphere Watch (GAW) of the World Meteorological Organization (WMO, 2007a), regional programmes such as the European Monitoring and Evaluation Programme (EMEP), and national air pollution monitoring networks.

In Europe the measurement capacity for VOCs in the atmosphere is diverse. On the one hand, several laboratories maintain long-term and high-quality measurements based on sophisticated quality assurance/quality control (QA/QC) systems, high-quality standard gases, previous intercomparison activities, and audits by the World Calibration Centre for VOCs. On the other hand, the performance of other laboratories is limited by a lack of such quality measures and by the fact that there are no commonly agreed-upon guidelines concerning standards, homogenised quality assurance procedures and measurement methods. In Table 1 of the WMO GAW Report No. 171 (WMO, 2007b) 17 priority VOCs (NMHCs and oxygenated VOCs) were identified and general quality assurance recommendations were defined (Table 2). The European infrastructure network project ACTRIS (Aerosols, Clouds, and Trace gases Research InfraStructure) has expanded the priority substances to further NMHCs described in Table 1. Furthermore, measurement guidelines and a quality management system were developed under ACTRIS to harmonise trace gas measurements of NMHCs in Europe (http://www.actris.net/Project/WorkPackages/WP4/tabid/4428/Default.aspx). One objective of ACTRIS was to assess the current NMHC measurement capacity in Europe and to investigate the analytical performance of laboratories in terms of data quality objectives (DQOs) for repeatability and uncertainty. Strict DQOs were defined in ACTRIS (Table 2) and were adopted by the GAW Scientific Advisory Group for Reactive Gases during their meeting in October 2014. Whilst in the WMO GAW Report No. 171 DQOs are defined for accuracy and precision, these have been replaced in ACTRIS with uncertainty (in the sense of expanded combined uncertainty with coverage factor $k = 2$; JCGM, 2008) and repeatability (which characterises the short-term standard variation in multiple measurements).

VOC species are normally measured with gas chromatography coupled to either a flame ionisation detector (GC-FID) or a mass spectrometer (GC-MS). Furthermore, proton transfer reaction mass spectrometry (PTR-MS) is also used for the measurement of oxygenated VOCs, terpenoids, dialkenes, and aromatics. While PTR-MS analyses VOCs from air samples directly, GC-based techniques need a preconcentration step. Here VOCs are either analysed immediately after sampling onto suitable adsorbents (online) or they are collected in specially treated steel or glass cylinders or on cartridges filled with adsorbents and analysed later in the laboratory (offline). Problems which can occur are chemical reactions in the samples (due to e.g. reactions with ozone), adsorptive losses, memory effects or leaks, losses during the preconcentration and the desorption steps, chemical reactions during thermal desorption, insufficient separation on the chromatographic column and misidentification, peak overlap, and inaccurate quantification (Helmig, 1997, 1999; Helmig and Vierling, 1995; Koppmann et al., 1995; Parrish and Fehsen-

Table 1. Assigned mole fractions (error-weighted means with expanded uncertainties) for NMHC_N_2 and NMHC_air (nmol mol^{-1}).

	Error-weighted mean ± expanded uncertainty (nmol mol^{-1})			
	NMHC_N_2		NMHC_air	
	cylinder 1	cylinder 2	cylinder 1	cylinder 2
Alkanes				
ethane	1.071 ± 0.016	1.118 ± 0.016	1.871 ± 0.037	1.904 ± 0.041
propane	1.061 ± 0.014	1.104 ± 0.015	1.608 ± 0.025	1.611 ± 0.023
n-butane	1.025 ± 0.028	1.076 ± 0.015	1.407 ± 0.019	1.407 ± 0.015
methylpropane	1.051 ± 0.011	1.114 ± 0.013	0.778 ± 0.026	0.765 ± 0.024
n-pentane	1.031 ± 0.012	1.092 ± 0.017	0.834 ± 0.012	0.834 ± 0.014
2-methylbutane	1.011 ± 0.011	1.075 ± 0.014	1.669 ± 0.029	1.654 ± 0.028
n-hexane	1.019 ± 0.013	1.083 ± 0.014	0.157 ± 0.006	0.151 ± 0.006
2-methylpentane	1.025 ± 0.014	1.089 ± 0.014	0.343 ± 0.025	0.348 ± 0.021
3-methylpentane			0.195 ± 0.009	0.194 ± 0.008
2,2-dimethylbutane			0.257 ± 0.038	0.256 ± 0.033
2,3-dimethylbutane			0.070 ± 0.020	0.072 ± 0.020
cyclohexane			0.140 ± 0.005[a]	0.141 ± 0.009[a]
n-heptane	1.011 ± 0.011	1.077 ± 0.012	0.443 ± 0.008	0.463 ± 0.010
n-octane	1.011 ± 0.021	1.076 ± 0.023	0.443 ± 0.008[a]	0.463 ± 0.010[a]
2,2,4-trimethylpentane	1.028 ± 0.012	1.095 ± 0.012	0.145 ± 0.008[a]	0.144 ± 0.008[a]
Alkenes				
ethene	1.065 ± 0.015	1.127 ± 0.015	2.531 ± 0.034	2.532 ± 0.035
propene	1.030 ± 0.013	1.091 ± 0.016	0.571 ± 0.020	0.552 ± 0.014
1-butene	1.007 ± 0.031	1.070 ± 0.033	0.114 ± 0.006[a]	0.109 ± 0.006[a]
2-methylpropene			0.858 ± 0.039[a]	1.081 ± 0.049[a]
trans-2-butene	1.024 ± 0.018	1.088 ± 0.019	0.074 ± 0.003[a]	0.075 ± 0.003[a]
cis-2-butene	1.008 ± 0.011	1.069 ± 0.013	0.066 ± 0.002	0.067 ± 0.002
1,3-butadiene[b]	1.024 ± 0.025	1.087 ± 0.024	0.066 ± 0.011	0.062 ± 0.014
1-pentene	1.001 ± 0.012	1.086 ± 0.036	0.048 ± 0.014[a]	0.044 ± 0.013[a]
trans-2-pentene	0.984 ± 0.015	1.042 ± 0.018	0.057 ± 0.004	0.058 ± 0.004
cis-2-penten			0.033 ± 0.003	0.032 ± 0.003
2-methyl-2-butene			0.125 ± 0.020	0.121 ± 0.006
isoprene[b]	2.039 ± 0.038	2.178 ± 0.034	0.021 ± 0.008	0.022 ± 0.006
Alkynes				
ethyne	1.020 ± 0.026	1.118 ± 0.024	1.467 ± 0.032	1.485 ± 0.039
propyne			0.065 ± 0.019[a]	0.065 ± 0.017[a]
Aromatic compounds				
benzene	1.022 ± 0.012	1.091 ± 0.013	0.460 ± 0.006	0.458 ± 0.007
toluene	1.021 ± 0.048	1.222 ± 0.039	1.709 ± 0.059	1.737 ± 0.055
ethylbenzene	1.017 ± 0.057	1.182 ± 0.057	0.245 ± 0.010	0.247 ± 0.008
m,*p*-xylene	2.035 ± 0.117	2.569 ± 0.108	0.884 ± 0.038	0.882 ± 0.036
o-xylene	1.047 ± 0.097	1.180 ± 0.095	0.279 ± 0.019	0.282 ± 0.023

[a] Assigned mole fractions were determined only with results from HPB; [b] arithmetic mean of measurements instead of error-weighted mean. Compounds: priority VOCs as defined in the WMO GAW Report No. 171 (WMO, 2007b). Further WMO GAW priority VOCs which were not investigated here: monoterpenes, dimethylsulfide, formaldehyde, methanol, ethanol, acetone, acetonitrile.

feld, 2000; Plass-Duelmer et al., 2002; Rudolph, 1999; Westberg and Zimmerman, 1993).

Several NMHC intercomparisons have been carried out in the past on European and global scales. Generally, these aimed at an evaluation of the quality of VOC measurements without data quality objectives defining a threshold to differentiate between higher- and lower-quality data (e.g. NOMHICE: Apel et al., 1994, 1999, 2003; AMOHA:

Table 2. ACTRIS and former GAW[1] data quality objectives (DQOs). Numbers express the expanded uncertainty (coverage factor $k = 2$), and the repeatability (standard deviation).

	GAW[1] uncertainty	GAW[1] repeatability	ACTRIS[2] uncertainty	ACTRIS[2] repeatability
alkanes	10 %	5 %	5 %	2 %
alkenes incl. isoprene	20 %	15 %	5 %	2 %
alkynes	15 %	5 %	5 %	2 %
aromatics	15 %	10 %	5 %	2 %
mole fraction $< 0.1\,\mathrm{nmol\,mol^{-1}}$ (ppb)	0.02 ppb	0.015 ppb	0.005 ppb	0.002 ppb

[1] From WMO GAW Report No. 171 (WMO, 2007b); [2] GAW VOC Expert Group and the GAW Scientific Advisory Group for Reactive Gases during their meetings in Daejong (South Korea, Oct 2014) as new GAW DQOs.

Plass-Duelmer et al., 2006; Slemr et al., 2002; GAW: Rappenglueck et al., 2006; Bernardo-Bricker et al., 1995; De Saeger and Tsani-Bazaca, 1993; Hahn, 1994; Pérez Ballesta et al., 2001; Romero, 1995; Volz-Thomas et al., 2002). NOMHICE (Nonmethane Hydrocarbon Intercomparison Experiment) and AMOHA (Accurate Measurements of Hydrocarbons in the Atmosphere) were two systematic multistage intercomparisons for NMHCs performed in North America and Europe, where the complexity of the NMHC measurements (numbers of compounds and sample gas mixtures) increased during the experiments. While in earlier intercomparisons the use of certified NMHC calibration standards was not common (Apel et al., 1994, 1999, 2003; De Saeger and Tsani-Bazaca, 1993; Hahn, 1994; Pérez Ballesta et al., 2001; Romero, 1995), multicomponent standards with certified NMHC mole fractions were circulated for analysis among the participating laboratories in more recent intercomparisons (Plass-Duelmer et al., 2006; Rappenglueck et al., 2006; Slemr et al., 2002). Within these studies the calibration with multicomponent NMHC calibration standards was superior to calibration with a single hydrocarbon species. Therefore, also in the ACTRIS intercomparison experiments all participating laboratories were asked to use certified multicomponent NMHC calibration standards, traceable to the GAW scale, for calibrating their instruments and for performing quality checks.

Eighteen stations or laboratories from nine European countries took part in this ACTRIS intercomparison exercise for the analysis of NMHCs. OVOCs were excluded due to their instability in pressurised cylinders at ambient mole fractions. Pressurised cylinders filled with NMHCs in nitrogen (in the following called NMHC_N_2) and NMHCs in whole air (in the following called NMHC_air) were analysed by the different laboratories using their own certified multicomponent NMHC calibration standard. The participants performed their measurements with GC-FID, GC-MS, or PTR-MS instrumentation. The performance of the different laboratories was examined with respect to compliance with the DQOs of ACTRIS and GAW (Table 2). Feedback was provided to the participants during a workshop, via analysis of technical details of each instrument, and the provision of recommendations for further characterisations and improvements.

This paper presents the findings of the intercomparison, with a focus on alkanes, alkenes, alkynes, and aromatic compounds. Results are used to discuss the status of current NMHC measurement capabilities in Europe 10–20 years after AMOHA, GAW, and NOMHICE intercomparisons and to discuss and evaluate issues with different instrumental setups used in the field.

2 Method section

2.1 Intercomparison approach

Eighteen European laboratories with 23 different GC instruments participated in this NMHC intercomparison exercise in 2012 (Fig. 1 with laboratory abbreviations, Tables S1–S2 in the Supplement). Additionally, two PTR-MS instruments analysed the NMHC mixtures (Table S2; results are shown only in the Supplement). It should be pointed out that the "PerkinElmer Online Ozone Precursor Analyzer" is the only commercially available all-in-one instrument tested in this study. All other instruments use combinations of commercially available parts and custom-built units.

The intercomparison exercise was performed in two loops (with nine participants each) in order to keep the total time for the exercise within a few months. All participants received two cylinders, one filled with NMHC_N_2 and one with NMHC_air.

2.2 Preparation of NMHC mixtures

The two NMHC mixtures, NMHC_N_2 and NMHC_air, were prepared in 10 L "Quantum" passivated aluminium cylinders (Air Products, purchased from National Physical Laboratory (NPL)). NMHC_N_2 was diluted with nitrogen (quality 5.0 from Linde AG, Germany) from a $\sim 100\,\mathrm{nmol\,mol^{-1}}$ uncertified mixture of 30 NHMCs (and several monoter-

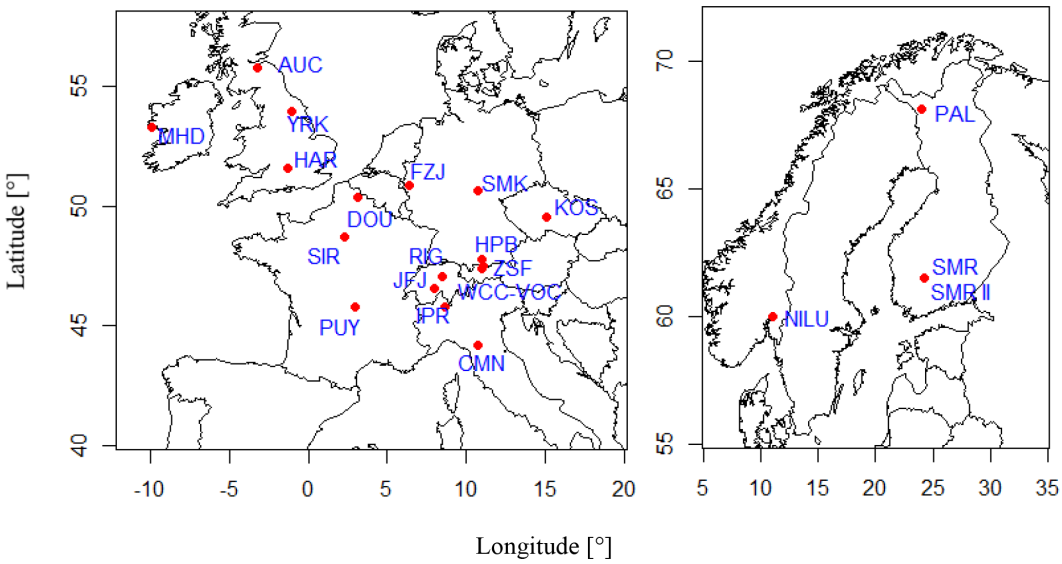

Figure 1. The participants of the ACTRIS NMHC intercomparison in Europe. Left: western, central, and southern Europe; right: northern Europe. AUC: Auchencorth, CMN: Monte Cimone; DOU: Mines des Douai; FZJ: Forschungszentrum Jülich; HAR: Harwell; HPB: Hohenpeissenberg; IPR: Ispra; JFJ: Jungfraujoch; KOS: Kosetice; MHD: Mace Head; NILU: Norwegian Institute for Air Research; PAL: Pallas; PUY: Puy-de-Dôme; RIG: Rigi; SIR: Sirta; SMR: GC-MS, and SMR II: PTR-MS (both at SMEAR II, Hyytiälä); SMK: Schmücke; WCC-VOC: Garmisch-Partenkirchen; YRK: York; ZSF: Zugspitze-Schneefernerhaus. For further details about the stations and instruments see Table S1–S2.

penes) in nitrogen (prepared by NPL for HPB on demand) into two cylinders by HPB. The resulting mole fractions in NMHC_N_2 were $\sim 1\,\mathrm{nmol\,mol^{-1}}$ (Table 1). The final pressure in the cylinders was ~ 120 bar. NMHC_air was filled with ambient air from Dübendorf (a suburban area of Zurich, Switzerland) in two 10 L cylinders, using a modified oil-free diving compressor (Model SA-6; RIX Industries, USA) on 31 October 2011. Due to the pressurisation, the water vapour condensed and the final humidity in the cylinders was very low (dew point $< -30\,°\mathrm{C}$, relative humidity $\sim 1\,\%$). The mole fractions of C_2–C_8 NMHCs in NMHC_air ranged from 0.03 to $2.5\,\mathrm{nmol\,mol^{-1}}$ (Table 1). The final pressure in the cylinders was ~ 80 bar. Mole fractions in NMHC_air were in the upper range of rural stations in Europe and higher than remote conditions (Helmig, 1997; Helmig et al., 2008; Read et al., 2009).

Three laboratories (WCC-VOC, HPB, and Empa) assigned NMHC mole fractions to the different cylinders before and after the intercomparison. Additionally, these two time-separated measurements were used to assess the stability of the NMHC mixtures. All three laboratories used certified NMHC calibration standards from the GAW Central Calibration Laboratory for NMHCs (NPL), which defines the scale for NMHC measurements in WMO GAW. The analytical systems of the three reference laboratories can be considered as sufficiently independent as different pre-concentration systems and chromatographic columns are used (see Table S2a–b).

Since HPB and Empa acted as reference laboratories, their data measured in the middle of the intercomparison were not used for the reference value determination. However for completeness, their values are displayed in Figs. 2–4 and S1–S4, and Tables S3–S6 together with those of the other participants and correspondingly marked as no "independent" results.

The NMHC mole fractions were usually assigned as error-weighted means (Barlow, 1989; Bronštejn, 2007) and are displayed with their corresponding expanded uncertainty (coverage factor $k = 2$, corresponding to 2σ or roughly a 95 % confidence interval) in Table 1. A more detailed description is given in the Supplement.

2.3 Measurement approach

A detailed measurement guideline was provided to the participants to ensure consistent and comparable measurements of the NMHC mixtures. All participants used the same provided pressure regulators (model 206A from Scott Specialty Gases, USA) and transfer lines (Silcosteel®, 1 / 16 in., ~ 2.5 m). The pressure regulator was mounted at least 24 h before the measurement onto the gas cylinder and connected to the transfer line. Afterwards, the regulator and the transfer line were flushed three times and an initial leak test was performed (observation of pressure drop during 10 min). The pressure regulator and the transfer line were kept pressurised for at least 24 h (with closed cylinder valve) for equilibration of surfaces. Additionally, this setup served as a static leak test.

All participants were asked to quantify the NMHC mole fractions using their own calibration standard (Table S2b) and to report their expanded measurement uncertainty (see Supplement). Within GAW and ACTRIS, the scale for NMHC measurements is defined by the Central Calibration Laboratory (CCL) NPL, which continuously compares their NMHC scale and the associated expanded combined uncertainty of typically 2 % with other NMI (National Metrology Institutes) (Grenfell et al., 2010). Though different scales might cause a bias of results by participants that are related to a non-NPL laboratory standard, this ACTRIS comparison study addresses the inter-laboratory compatibility related to the standardised GAW NMHC scale provided by NPL. The certified expanded uncertainties of the NPL of 2 % are generally much lower than deviations discussed in this paper, e.g. beyond the DQO of 5 %.

The composition and the mole fractions in the cylinders were unknown to all participants, except for the reference laboratories HPB and Empa (see above). The measurement procedure was the following: at least three calibration standard measurements, five measurements of NMHC_N_2, five measurements of NMHC_air, at least three calibration standard measurements, and a zero-gas measurement before and after the NMHC mixture measurements. Fourteen analyses were by GC-FIDs and nine by GC-MSs (Table S2). In this paper, results for 27 and 35 NMHCs are shown for NMHC_N_2 and NMHC_air, respectively. The three trimethylbenzenes and the monoterpenes present in NMHC_N_2 were not investigated in this intercomparison paper due to the lack of available data. The assigned NMHC mole fractions (with expanded uncertainties) are given in Table 1.

2.4 Data quality objectives (DQOs) for NMHC measurements

In the WMO GAW Report No. 171 (WMO, 2007b) general DQOs for different priority VOCs were defined (Table 2). Within the framework of ACTRIS, the list of priority compounds (Table 1) was expanded, and more challenging DQOs (ACTRIS DQOs) were defined (Table 2). Overall, ACTRIS DQOs are about a factor of 2 stricter than those in the GAW Report 171. The reason for the introduction of the ACTRIS DQOs was to detect trends of NMHCs more accurately, which currently decline by 1–8 % per year in Europe (Solberg, 2012, 2013, and references therein). These ACTRIS DQOs were also adopted by the GAW VOC Expert Group and the GAW Scientific Advisory Group for Reactive Gases during their meetings in Daejong (South Korea, October 2014). For the uncertainty, which describes the deviation from a reference value, the goals are set to 5 % for alkanes, alkenes (including isoprene), alkynes, and aromatics (and to 10 % for monoterpenes). Values express the expanded uncertainty with a coverage factor of $k = 2$. The goals in repeatability, defined as the standard deviation of the NMHC measurements, are 2 % for alkanes, alkenes (including isoprene), alkynes, and aromatics, and 5 % for monoterpenes. For mole fractions below 0.1 nmol mol^{-1} an absolute value of 0.005 nmol mol^{-1} is accepted as uncertainty, and 0.01 nmol mol^{-1} for monoterpenes.

In the results section the measurement performance is compared against these DQOs by ACTRIS and adopted by GAW (Table 2). Hereby the uncertainties u_{ref} of the assigned reference values (Table 1) need to be taken into account. Thus, a result fulfils the ACTRIS DQO if the deviation from the reference is less than the 5 % deviation class defined as

$$5\%\ \text{class} = \sqrt{\mathrm{DQO}^2_{\mathrm{ACTRIS}} + u^2_{\mathrm{ref}}}. \quad (1)$$

For the 10 % deviation class (10 % class), the respective former GAW DQO (Table 2) is applied.

2.5 C response for GC-FID systems

A GC-FID system can be characterised for losses or artefacts by making use of the known carbon response, the so-called C response (Plass-Duelmer et al., 2002). When the C responses for the various NMHC compounds are calculated, they should agree within a few percent, except for ethyne (Burns et al., 1983; Dietz, 1967; Faiola et al., 2012; Gong and Demerjian, 1995; Scanlon and Willis, 1985; Sternberg et al., 1962). The C response R_i for each compound i was calculated as follows:

$$R_i = \frac{A_i^{\mathrm{std}} - A_i^{\mathrm{b}}}{m_i^{\mathrm{std}} N_i V^{\mathrm{std}}}, \quad (2)$$

where A_i^{std} and A_i^{b} are the peak areas of compound i in the calibration standard (std) and the blank (b), respectively; m_i^{std} denotes the certified mole fraction of the calibration standard; N_i the number of C atoms in compound i; and V^{std} the sampled volume of the calibration standard.

When comparing the C response values in the calibration standard and in NMHC_N_2, the C responses should ideally be identical. Deviation points towards either artefacts in the analytical system (e.g. breakthrough during trapping, adsorptive losses, peak overlap, changes on active sites) or in the FID due to sample matrix effects influencing the flame. For easier comparison, the C responses were normalised by the average C response of the available C_4–C_6 alkanes (highlighted in yellow in Fig. 4). As some stations did not report C_2–C_3 alkanes (e.g. HPB_B, FZJ_A) and additionally breakthrough in C_2 compounds could have occurred, only C_4–C_6 alkanes were taken into account. For two-column systems, the average C response of the second column was determined using C_7–C_8 alkanes, benzene, and toluene (highlighted in green in Fig. 4). Any individual C response deviating more than 10 % from the average C response was not considered in the normalisation process.

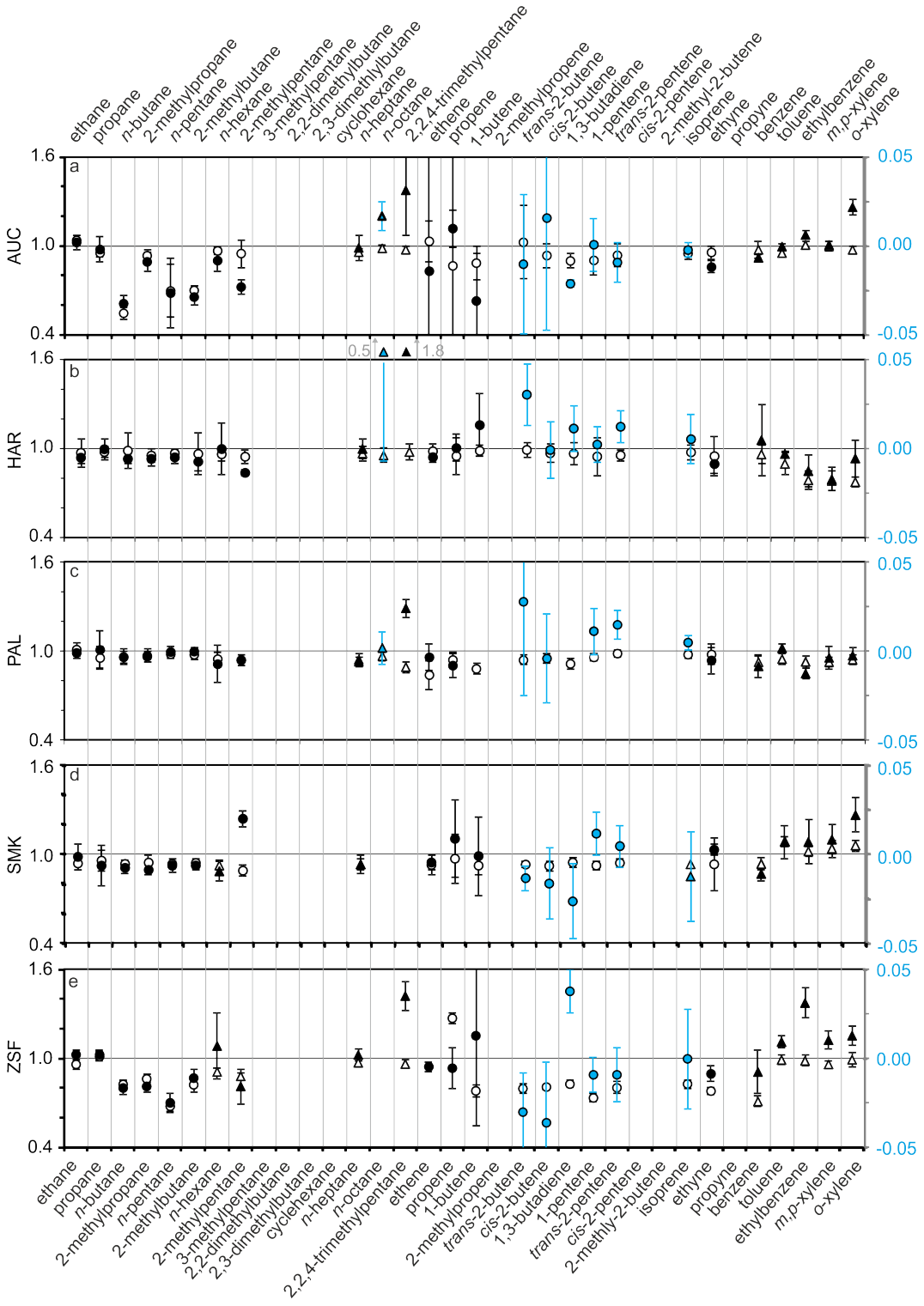

Figure 2. Mole fractions for NMHC_N_2 and NMHC_air normalised to the assigned values. Circles (O) indicate separation column one, triangles (Δ) separation column two. Open symbols indicate NMHC_N_2, filled symbols NMHC_air. Black symbols indicate results for mixing ratios > 0.1 nmol mol^{-1} (left axis, ratios to assigned values), blue symbols for < 0.1 nmol mol^{-1} (right axis, difference to assigned values, see text). Error bars show the expanded combined uncertainties.

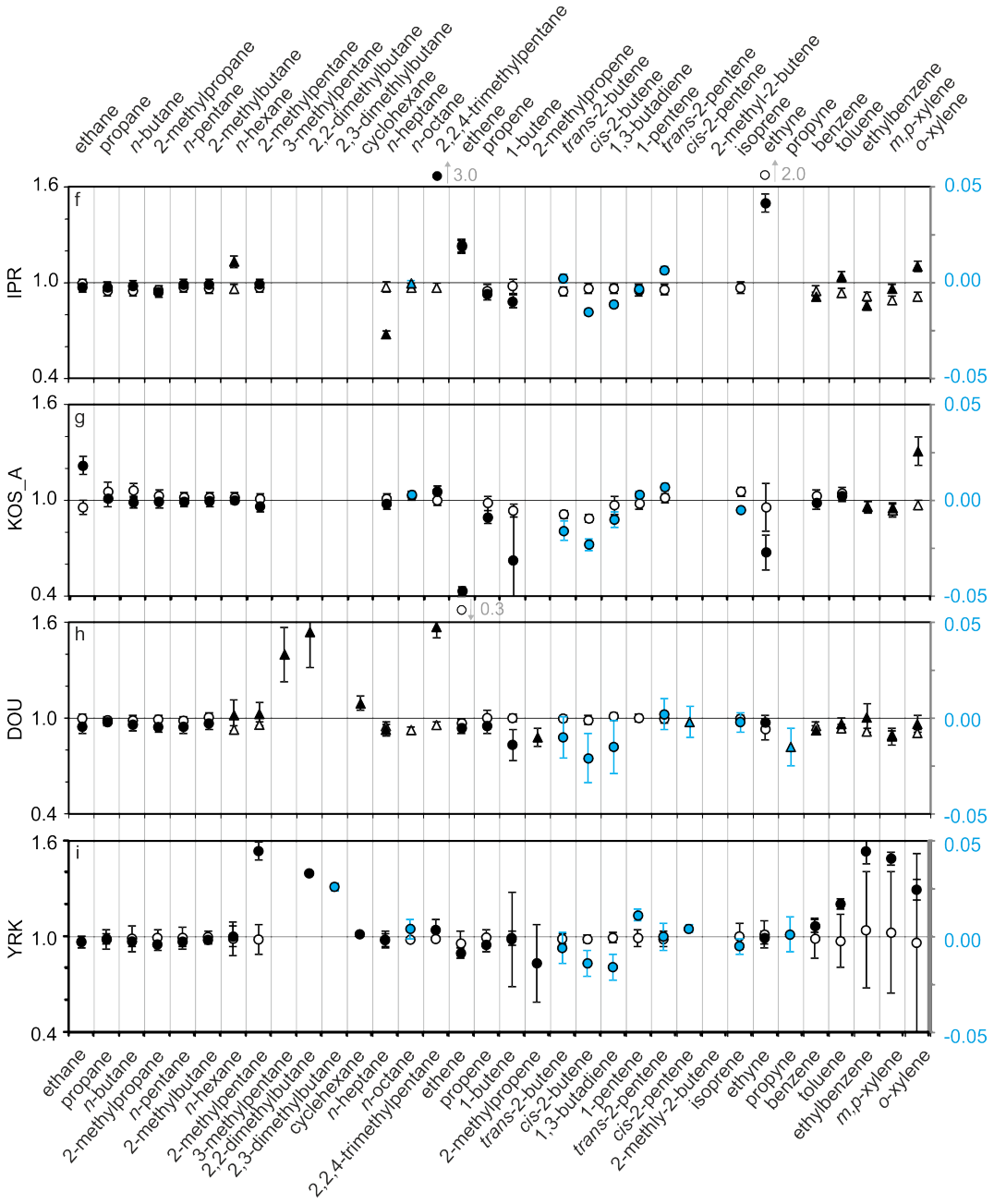

Figure 2. Continued.

3 Results and discussion

3.1 Overview of results

Mole fractions for NMHC_N_2 and NMHC_air normalised to the assigned values (Table 1) are shown for each participant and compound in Fig. 2 (black symbols, left *y* axis). For compounds with mole fractions $< 0.1\,\mathrm{nmol\,mol^{-1}}$ the difference in $\mathrm{nmol\,mol^{-1}}$ to the assigned values is shown (blue scale on the right *y* axis). Error bars represent the total expanded uncertainty as stated by the participant. Box-and-whisker plots for all results are presented in Fig. 3. Results are in compliance with the ACTRIS DQO if they fall into the 5 % class (Eq. 1). For compounds with mole fractions below $0.1\,\mathrm{nmol\,mol^{-1}}$, this class reflects a deviation of $5\,\mathrm{nmol\,mol^{-1}}$ plus uncertainty of reference. In addition, for GC-FID systems C responses were calculated and are depicted in Fig. 4. More detailed information on the performance of different measurement systems is given in Tables S3–S6 as well as Figs. S1–S4 in the Supplement.

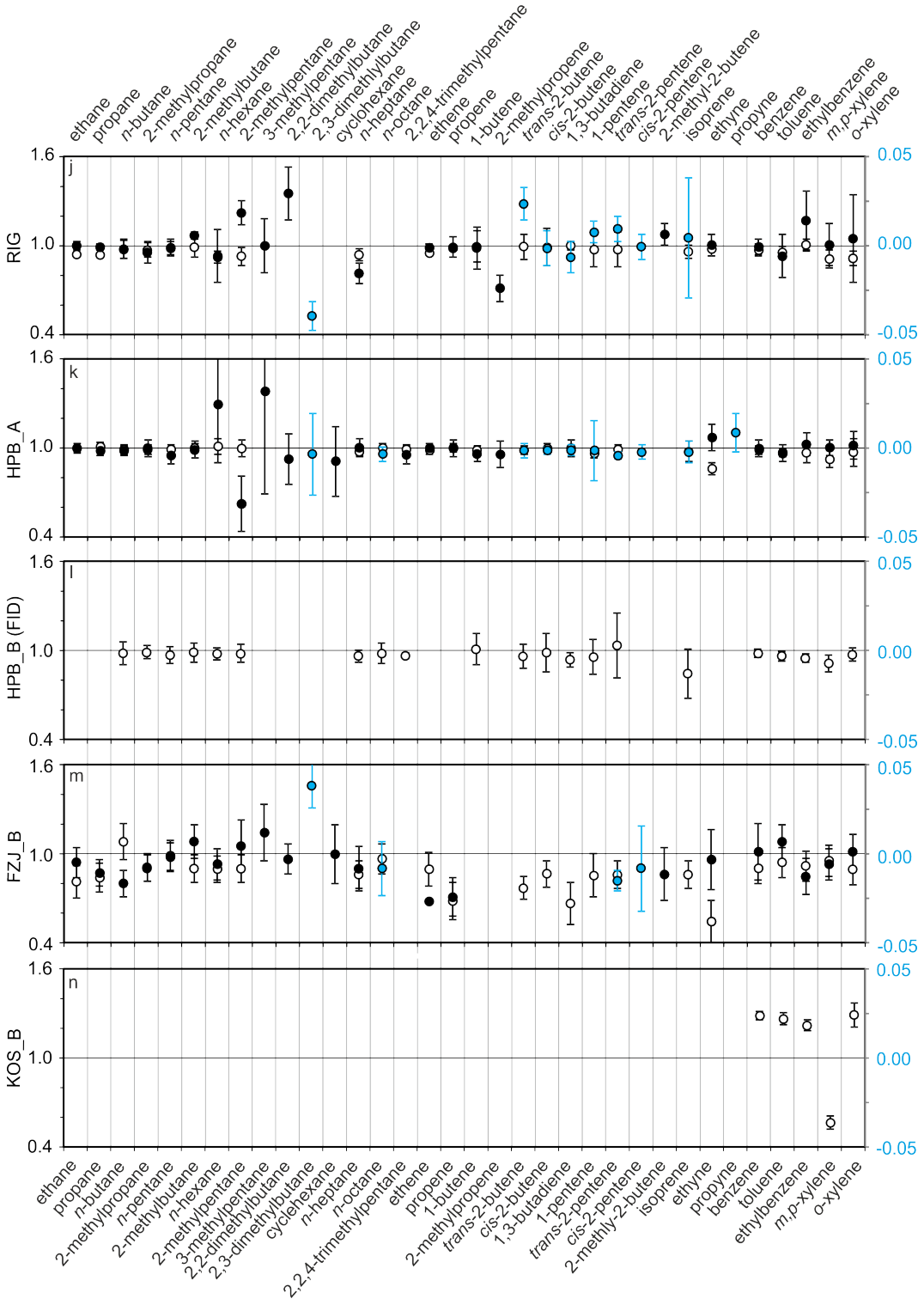

Figure 2. Continued.

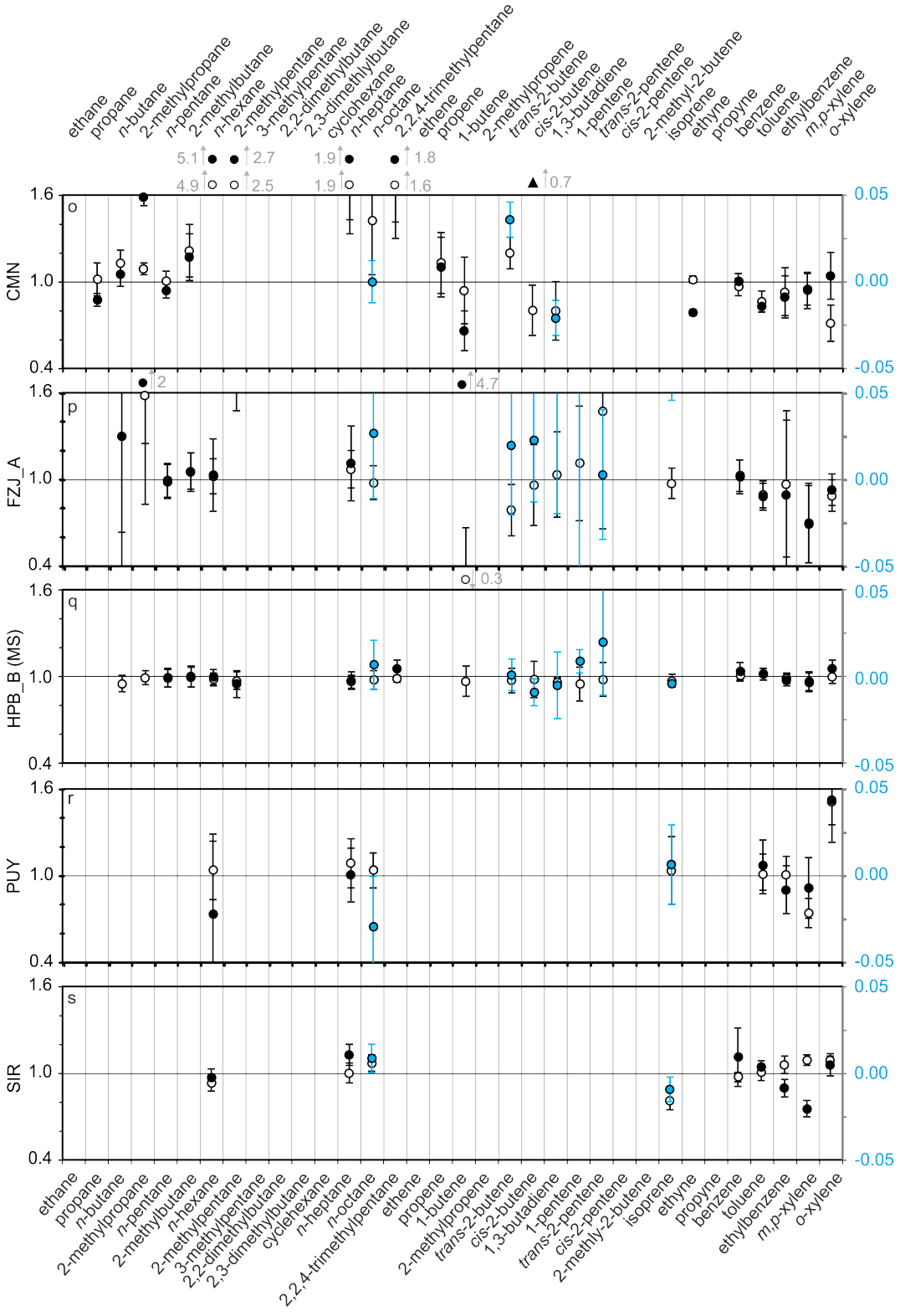

Figure 2. Continued.

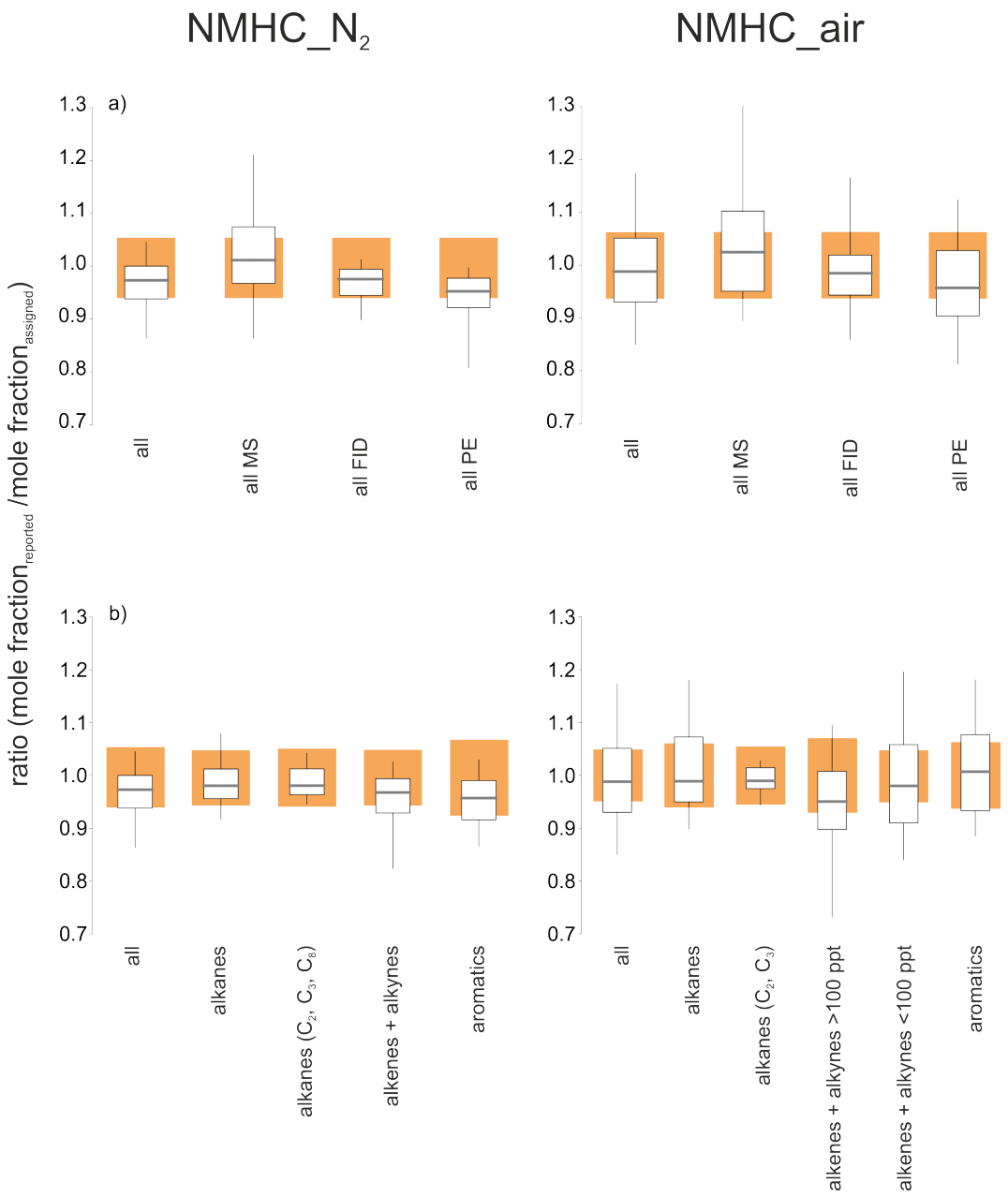

Figure 3. Box plots for NMHC_N_2 (left) and NMHC_air (right) relative to the assigned mole fractions (Table 1). **(a)** Overview box plot for all results, **(b)** box plot for the different compound classes, **(c)** box plot for alkanes, **(d)** box plot for alkenes and alkynes, **(e)** box plot for aromatics. The white box stretches from the 25th percentile to the 75th percentile, containing the median in between. The whiskers end at the 10th and 90th percentile. The orange box indicates the 5 % class (see Table 2 and Eq. 1). MS: mass spectrometer; FID: flame ionisation detector; PE: PerkinElmer.

For the $\sim 1\,\mathrm{nmol\,mol^{-1}}$ in nitrogen (NMHC_N_2) nearly 62 % of all results were within the 5 % class (ACTRIS DQOs), and nearly 90 % within the 10 % class (former GAW DQOs; Table 2) (Fig. 3 and Tables S3–S4). The best performance was achieved for alkanes with 65 % of the submitted alkane data within 5 % of the reference; for the subgroup of C_2–C_3 alkanes even 81 % of the submitted data were within the 5 % class (Table S3, Fig. 3b). C_4–C_7 alkanes were more challenging, and more deviations to the assigned mole fractions were observed (Fig. 2). For alkenes and aromatic compounds the percentages of results within the 5 % class were 58 and 47 %, respectively (Table S3, Fig. 3). Results for aromatics reveal a distinct tendency to be underestimated in NMHC_N_2 (Fig. 3e). Nevertheless, 63 % of the results match the 5 % class (Table S3), reflecting the rather large uncertain-

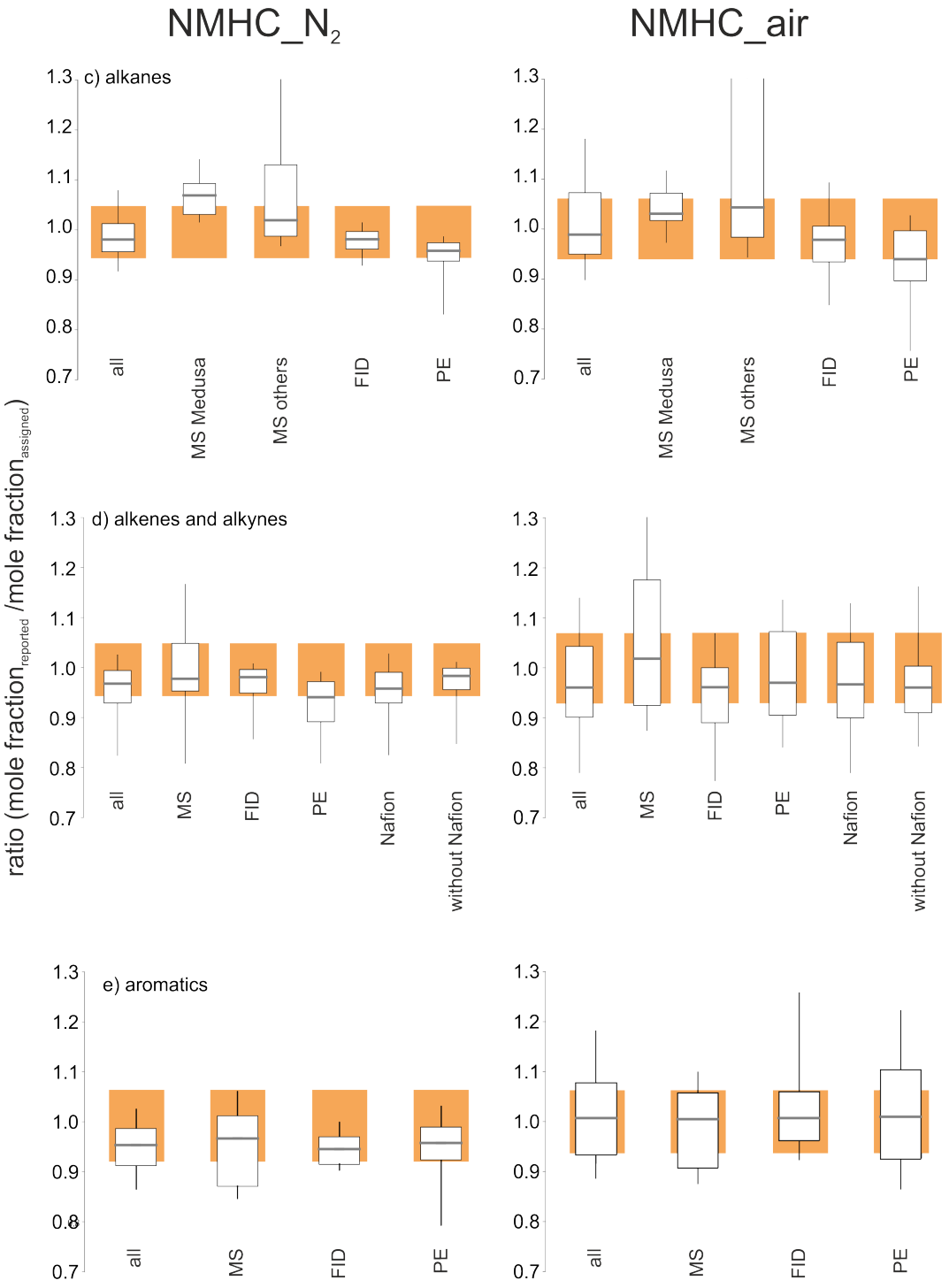

Figure 3. Continued.

ties of some reference values for aromatics (e.g. *o*-xylene, Table 1).

For NMHC_air, generally more frequent and larger deviations from the assigned values were observed compared to NMHC_N_2, and 50 % of all results were found within the 5 % class (79 % within the 10 % class) (Fig. 3, Tables S3 and S4). Only C_2–C_3 alkanes reveal a rather good performance, with 84 % of the respective data within the 5 % class. Compared to NMHC_N_2, the tendency to underestimate aromatic compounds is not observed (Fig. 3e).

The repeatability of the instruments was evaluated as the standard deviation (1σ) of the five measurements for both

NMHC mixtures. The majority of the participants submitted a relative repeatability in NMHC_N_2 within the former GAW DQO (± 5 % for alkanes and alkynes, ± 10 % for aromatics and ± 15 % for alkenes including isoprene), 70 % even within ± 2 % (ACTRIS DQO), independent of the detector type. Poor repeatability was mostly linked with poor chromatographic resolution (see Tables S5 and S6).

In the following, reasons for deviations larger than the stated quality objectives will be discussed.

3.2 Uncertainty estimations of the NMHC measurements

Performing a complete uncertainty estimation is critical to the quality of the data. Nevertheless, only the participants DOU, KOS (both systems), RIG, HPB (both systems), JFJ, MHD, NILU and ZSF provided a thorough analysis (see "Determination of assigned values (error-weighted means) for NMHC mixtures" in the Supplement) of their expanded uncertainties (error bars in Fig. 2). All other participants calculated their measurement uncertainties only partially (e.g. only reporting repeatability). Generally, for many results the uncertainties were underestimated and, even combined with the uncertainties of the reference values, do not comprise the deviation from the assigned values. Thirty-six percent of results in NMHC_N_2 were out of the stated uncertainty ranges, and 35 % in NMHC_air. As the expanded uncertainty corresponds to the 95 % confidence interval, it would be expected that not more than 5 % of the results deviate by more than the uncertainty from the assigned values. This needs to be improved in programmes like GAW and EMEP, as realistic uncertainty estimation is essential for the user, e.g. in model validation.

Critical in this evaluation are the assigned values; if these are biased relative to the "true" values, deviations may occur. However, in NMHC_N_2 there was a check by a common dilution factor relative to a NPL-certified standard of identical relative composition, which strongly supports the determined mole fractions within better than 2 % and does not indicate any bias. For NMHC_air, we rely on the uncertainty evaluation of the reference values by the reference laboratories, which is considered a realistic estimate. Though the assigned values are generally higher than the majority of the participants' results (Fig. 3), they are typically between the median and the 75-percentile or 90-percentile values with partially contradicting deviations for the various techniques; e.g. alkanes derived from MS are high, whereas those from FID are low compared to the reference (Fig. 3c). Furthermore, deviations in participants' results are similar for NMHC_N_2 and NMHC_air (e.g. Fig. 3a, various results in Fig. 2), supporting the assigned values in NMHC_air based on reliable NMHC_N_2 determination (see below).

3.3 Calibration procedure

One essential step on the way to high-quality NMHC data is the use of an adequate calibration procedure. The participants calibrated their NMHC measurements either directly against certified multicomponent standards (one-step calibration) or against whole air working standards, which in turn are related to a certified multicomponent standard (two-step calibration done by CMN and Medusa systems). The systems using a NPL (the GAW Central Calibration Laboratory for NMHCs) standard for direct, one-step calibrations (Table S2b) generally exhibited a good performance. Since the NMHC_N_2 mixture and the NPL calibration standard virtually comprise the same matrix, complexity, and manufacturer, observed deviations for sites referring to the NPL scale should be within the repeatability of the instruments. This is not the case for some participants and compounds, and it points to unidentified sample transfer issues. The mole fraction range of the used NPL standards (e.g. 2, 4, or 10 nmol mol^{-1}) and date of production apparently did not affect the quality of the results (Fig. 2, Table S2b).

The systems FZJ_B, FZJ_A, MHD, and PUY used different certified NMHC calibration standards (Table S2b). If a systematic offset between different scales exists, it should result in systematic deviations from the assigned values. FZJ_B, FZJ_A, and MHD all used calibration standards from Apel-Riemer, but the observed deviations from the assigned values are random (e.g. deviations for alkanes are of different extent and sign (Fig. 2m, p, and v). Obviously other instrumental issues (e.g. chromatographic resolution, non-linearity of MS detector) affected these results and therefore systematic differences between the different calibration scales cannot be assessed.

The Medusa instruments (JFJ, MHD, and NILU) generally overestimated the NMHC mole fractions (Figs. 2u–w and 3b). However, the excellent repeatability suggests that the systems run much better than the deviations indicate. Thus, a significant issue might arise from the fact that Medusa instruments and CMN are calibrated with whole air working standards using a two-step calibration. Direct calibration by certified NMHC standards appears to be superior to whole air working standards for NMHCs.

3.4 GC-FID systems

In order to analyse the performance of the GC-FID systems, the normalised C response factors for the calibration standards and NMHC_N_2 were compared (Fig. 4). Though identical C- responses are expected, several GC-FID systems tend to slightly underestimate NMHCs in NMHC_N_2 compared to the calibration standard (Figs. 4 and 2). Even more surprising was the fact that in some of the systems which have two separation columns, a lower normalised C response for NMHC_N_2 compared to the calibration standard was observed in only one column, e.g. AUC (in the PLOT column)

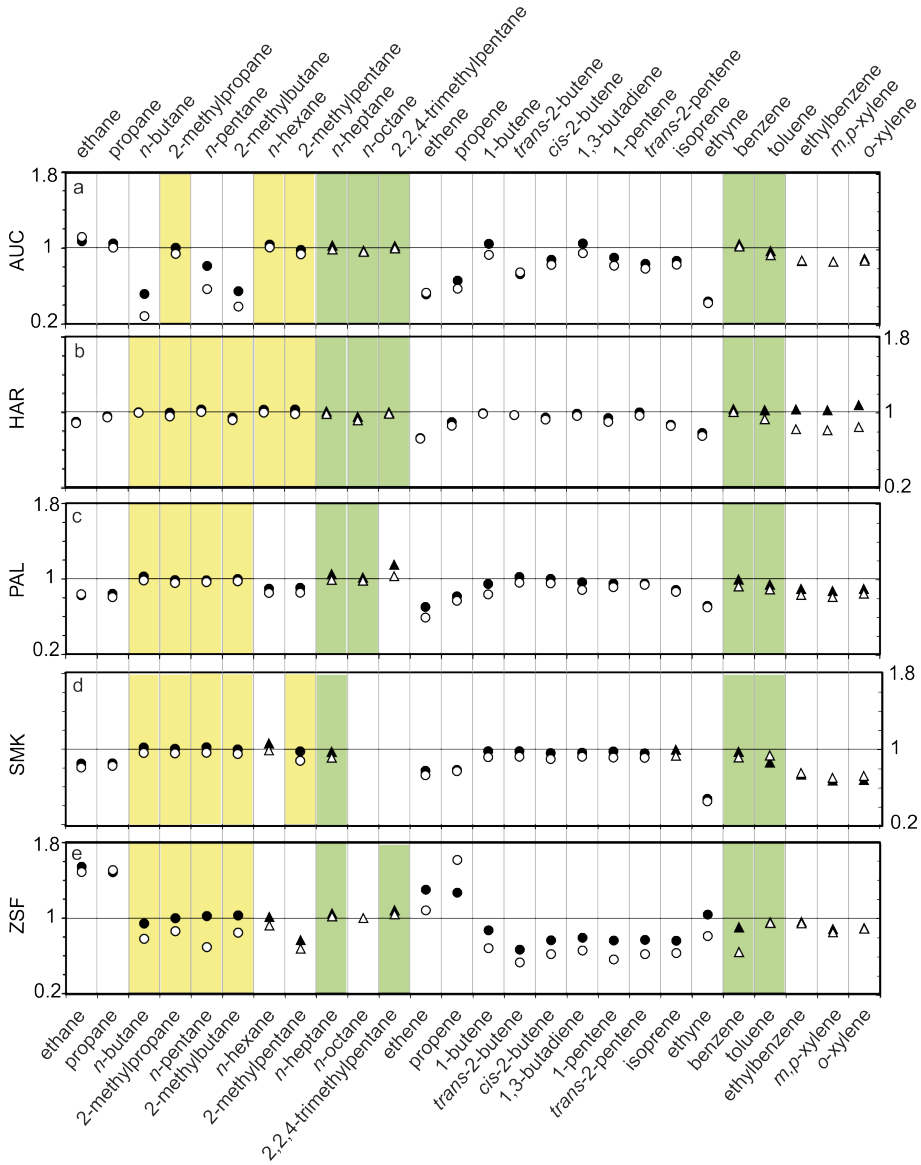

Figure 4. FID C responses as an indicator of NMHC measuring system artefacts. Circles (O) indicate column one, triangles (Δ) column two. Filled symbols indicate the C responses in the calibration standard, open symbols the C responses in NMHC_N_2. The C responses for column one were normalised by the average C response of the available C_4–C_6 alkanes (highlighted in yellow), column two by the average C response of the available C_7–C_8 alkanes, benzene, and toluene (highlighted in green). If an individual C response deviated by more than 10 % from the average C response, the value was not considered in the normalisation process.

and DOU (in the CP Sil-5 CB column, Table S2b) (Fig. 4a, h). The latter excluded sample transfer problems from the cylinder to the GC but pointed towards systematically different carrier gas or detector sensitivity conditions between analyses of calibration standard and NMHC_N_2. Overall, these discrepancies cannot be explained as general features but must be related to individual technical issues of the respective GC systems.

In general, there are several potential reasons for the deviations of C response factors between the calibration standard and the NMHC_N_2 and from the expected value of 1. They include losses of sample due to breakthrough, incomplete desorption or losses on walls, poor chromatographic resolution with inadequate peak separation or shape, and other artefacts (e.g. water management) and are addressed below.

3.4.1 Problems with C_2–C_3 hydrocarbons

More than 80 % of the C_2–C_3 hydrocarbons were reported within the 5 % class in NMHC_N_2 (Tables S3–S4, Figs. 2 and 3b). However, a few systems reported C_2–C_3 hydrocarbon values even outside the range of ± 10 %. The systems of

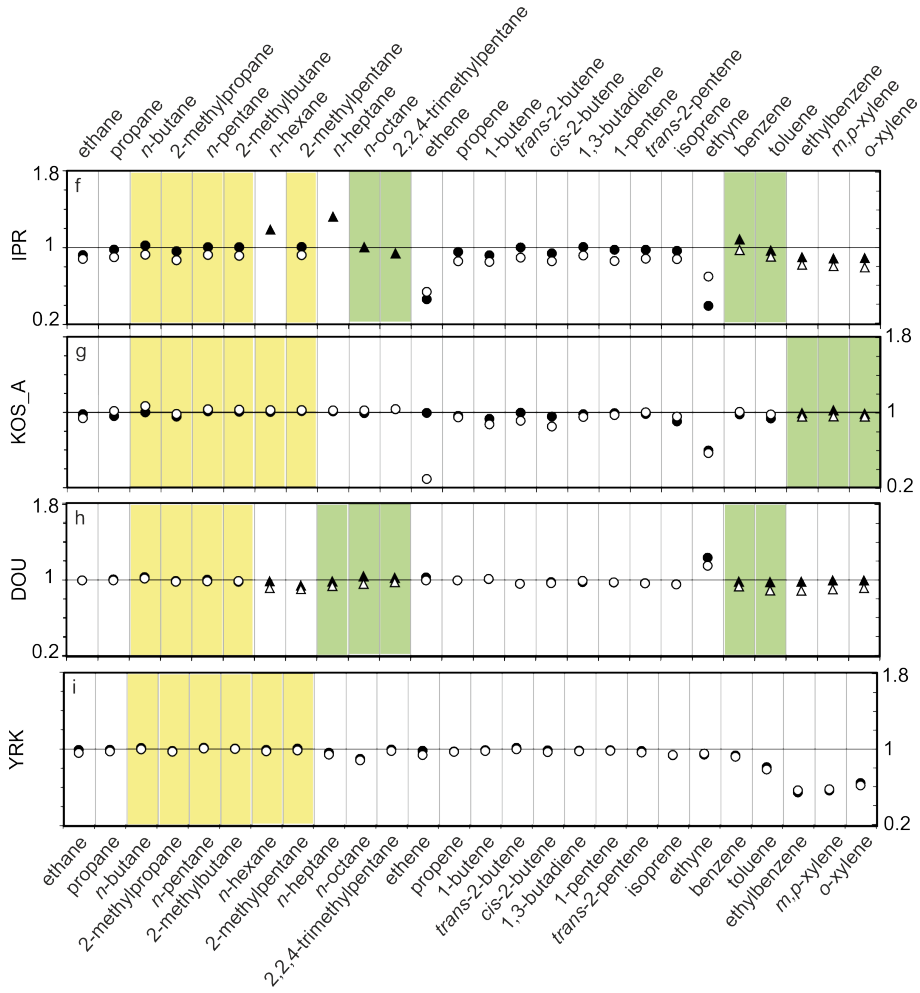

Figure 4. Continued.

AUC (only for ethene and propene), HAR, PAL, SMK, and IPR showed losses of C_2–C_3 hydrocarbons in the normalised C response plots, and ZSF and AUC (only for ethane) gave enhanced values (Fig. 4a–f). All of these six mentioned systems used a PerkinElmer TurboMatrix, which contains an air-toxics/air-monitoring trap, applies a 2 mL min^{-1} outlet split between trap and column, and has a two-column configuration with a Deans switch (Table S2). Sample breakthrough could be a specific issue of the PerkinElmer trap for these compounds. Badol et al. (2004) reported breakthrough for ethene and ethyne when the sampling volume exceeded 900 mL. However, no systematic influence of sample volume (up to 1360 mL) and trapping temperature (−40 or −30 °C) could be identified (Table S2). For example IPR used −30 °C as the adsorption temperature and a very large sample volume (1360 mL) but showed only moderate loss of C_2 hydrocarbons compared to e.g. SMK with 500 mL sample volume and larger losses (Fig. 4). For ZSF, the C responses for the C_2–C_3 hydrocarbons were even enhanced (> 1) compared to all other hydrocarbons (Fig. 4e). This system was affected by the displacement to 2650 m a.s.l shortly before this intercomparison exercise. Thus, most probably changed pressure and flow rates caused deviations during thermal desorption and might have affected the chromatographic resolution and thus the measurement quality of several of the compounds investigated in this study. In Figs. 2e and 4e, distinct deviations are observed for many substances, including C_2–C_3 compounds (Fig. 4e). Another explanation for decreased C responses for C_2–C_3 hydrocarbons could be a split issue during column injection. If a pressure pulse builds up in the thermodesorption phase, the split ratio might vary during the injection period causing different split ratios for high- and low-volatility hydrocarbons. Further, this pressure pulse could potentially influence the Deans switch. However, the systems of DOU and KOS_A also have Deans switches (with different thermodesorbers, Markes and Entech, respectively) and did not show losses of C_2–C_3 hydrocarbons. Based on these results it is not possible to distinguish between split, and trap issues and this needs further investigation.

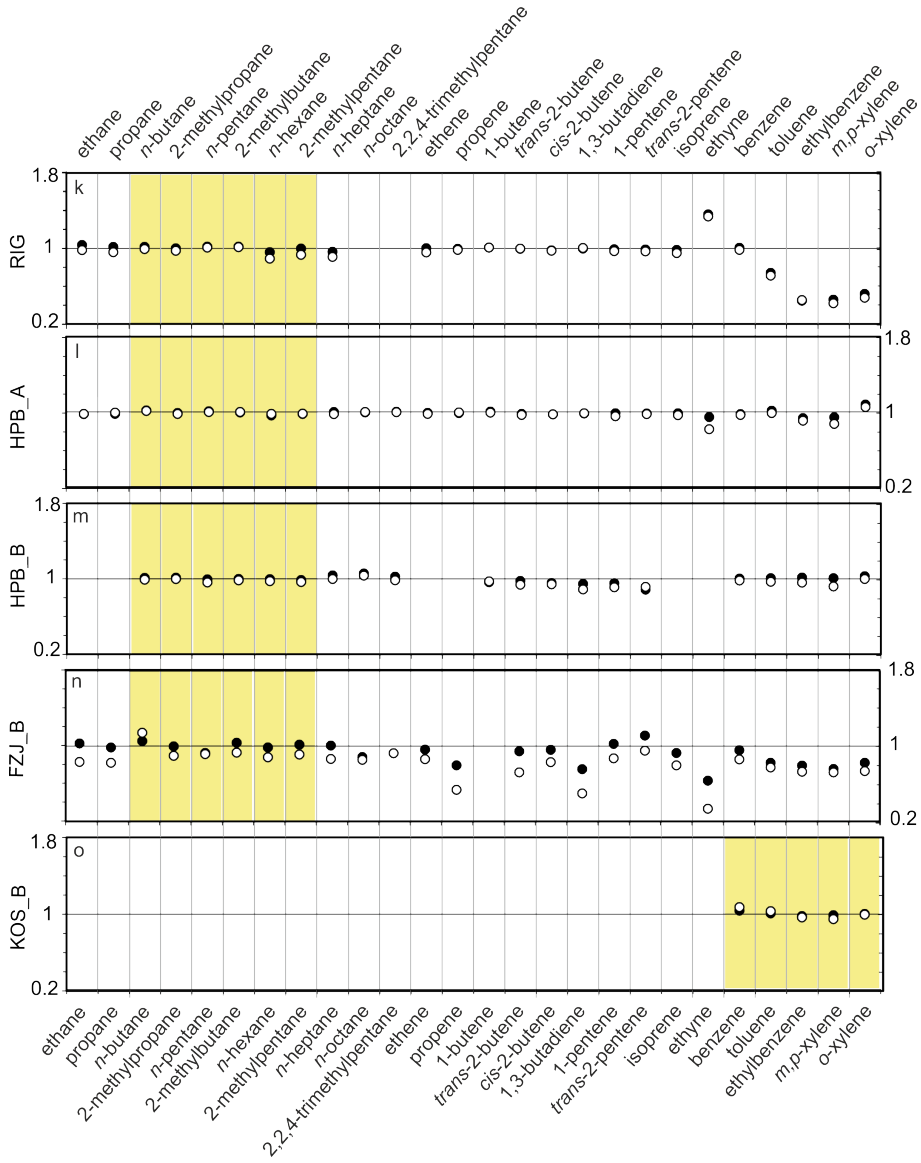

Figure 4. Continued.

In NMHC_air the results of the low-boiling alkanes (up to C_5) were more scattered compared to NMHC_N_2 (Figs. 2 and 3) mainly due to limits in chromatographic resolution (see Sect. 3.4.5). Despite apparent losses in specific systems (C response < 1, Fig. 4), most systems did not show deviations for the C_2–C_5 alkanes. However, the reference laboratory WCC-VOC observed specific matrix problems in NMHC_air in their GC-FID/FID system (not shown), which they attributed to extraordinarily high OVOC levels causing unusual artefacts in their trapping and desorption system. The high OVOC mole fractions of $\sim 200\,\mathrm{nmol\,mol^{-1}}$ were analysed by their PTR-MS system (see PTR-MS results in Supplement).

Low-boiling alkenes (C_2–C_5) showed partly substantial deviations to the assigned mole fractions in the AUC, PAL, SMK, ZSF, and IPR results, especially in NMHC_air (Figs. 3, 4a, c–f). Additionally to the aforementioned problems, alkene artefacts (see below) and, in the case of KOS_A, poor chromatography resolution contributed. The rather low mole fractions (< 100 ppt) did not affect the quality of the results (Fig. 3b).

3.4.2 Ethyne

For ethyne large differences in the C response factors (values between 0.3 and 1.4) were observed for the different stations (Fig. 4). Furthermore, large variations (up to 0.4) between the two C responses (calibration standard and NMHC_N_2) were evident. Based on the literature (Dietz, 1967; Scanlon and Willis, 1985; Sternberg et al., 1962) the effective car-

bon number is between 2 and 2.6, indicating a higher uncertainty of the C response for this compound. Thus, in the normalised C response figures ethyne is expected to be 1 or higher. This was actually observed for DOU, YRK, and RIG. Deviations between the laboratory standard and the ACTRIS NMHC_N_2 were observed at ZSF, DOU, HPB_A, and FZJ_B. Since at ZSF and FZJ_B observed deviations were not particular to ethyne but a general phenomenon for many compounds, both stations are not further considered in this specific discussion. The normalised C response of ethyne in the calibration standard of IPR was substantially lower than that of other stations (Fig. 4f).

Together with ethene, ethyne is the most difficult compound to be retained in air-toxics/air-monitoring traps (Badol et al., 2004). As AUC, HAR, PAL, SMK, ZSF, IPR, and KOS_A employ this type of traps, a breakthrough might be possible. However as already discussed, no conclusive behaviour, e.g. higher losses for higher sample volume and higher trapping temperature, was observed.

The instruments at DOU and HPB_A had in common that both employ an Al_2O_3/KCl PLOT column. However, other stations using the same type of column (YRK, RIG) did not show this feature. We are currently speculating about slightly different matrices between the calibration standard and NMHC_N_2 causing different interactions with active sites of the specific PLOT column, resulting in more or fewer losses.

Despite these losses observed in the C response factors, the difference to the assigned mole fractions were minor for six systems and moderate to substantial for 7 of 14 systems (larger than 10 % in either or both of the two NMHC mixtures) (Fig. 2), with often substantially different deviations for NMHC_N_2 and NMHC_air indicating matrix effects. This shows that it is essential to have ethyne in the calibration standard for direct calibration and that there is a need for thorough testing of matrix effects; e.g. real ambient air samples with higher humidity might result in higher breakthrough.

3.4.3 Alkene artefacts

Alkenes in NMHC_air exhibited largest differences to the assigned values (Fig. 2), especially pronounced for all systems which used Nafion® Dryer water traps, including PerkinElmer systems (Fig. 3d).

When using a Nafion® Dryer to remove humidity from the sample, potential artefacts in C_2–C_4 alkenes may occur depending on the status of the Nafion® Dryer (Gong and Demerjian, 1995; Plass-Duelmer et al., 2002, and references therein). Butene peaks (for 1-butene, *trans*-2-butene, and *cis*-butene) are frequently observed in zero-gas measurements due to Nafion® Dryer artefacts, and these blank values have to be subtracted in calibration or ambient air measurements. Instruments using a Nafion® Dryer reported blank values up to 0.35 nmol mol^{-1} for C_2–C_3 alkenes and up to 0.1 nmol mol^{-1} for C_4 alkenes. Combined with the fact that the mole fractions of C_4–C_5 alkenes were in the range of 0.02–0.12 nmol mol^{-1}, it is expected that substantial differences to the assigned values occur due to blank issues. For ethene and propene, however, such effects were comparably small due to much larger mole fractions up to 2.5 nmol mol^{-1} and blank values up to 0.25 nmol mol^{-1}. It should be noted that the samples measured here were not humid and thus the effects of water removal from the sample and the Nafion® Dryer behaviour cannot fully be assessed. Most participants were aware of the effects of a Nafion® Dryer and reported larger uncertainties of their values (Fig. 2).

3.4.4 Losses of aromatic compounds and C_6–C_8 alkanes

The C responses for the C_7–C_8 alkanes and for the aromatics were lower than 1 (Fig. 4), indicating losses in the analytical system. Lower C responses were observed either in both calibration standard and NMHC_N_2 (Fig. 4; AUC, PAL, SMK, IPR, YRK (except benzene), RIG, FZJ_B, and less evident in HPB_A) or only in NMHC_N_2 (Fig. 4; HAR, DOU, HPB_B, and FID). This effect was apparent in both intercomparison loops. This does not seem to be a general C response issue for aromatics, because in many systems not all aromatics showed a reduced C response (Fig. 4; KOS (both systems); for benzene: AUC, HAR, HPB (both systems), RIG, YRK) and several other systems showed only a reduced C response for NMHC_N_2 (Fig. 4; HAR, DOU, and HPB_B, FID). For these systems, systematic problems like insufficient desorption from the trap or adsorptive losses in the GC system can thus be excluded. However, adsorptive losses only in NMHC_N_2 might have occurred due to insufficient equilibration time and the flushing procedure of the respective pressure regulator and transfer lines. RIG reported lower C responses compared to the calibration standard for C_6–C_8 alkanes and aromatics (Fig. 4k). This was related to insufficient desorption temperature due to ice on the outer side of the Peltier-cooled trap which had built up during trapping.

In general, too-low desorption temperature from the trap can be excluded for the glass bead traps (70–130 °C, Table S2). For the air-toxics traps no losses of aromatics were observed for HAR (trap at 320 °C) (Fig. 4a). By contrast, losses prevailed at up to 380 °C (IPR), which were consequently not due to too-low desorption temperature (Fig. 4f). YRK results indicated losses which were not due to desorption temperature (Carbopack B and Carboxen 1000 at 350 °C) but were ascribed to adsorption on newly installed stainless-steel transfer lines. In the slightly more humid NMHC_air, YRK achieved relatively higher aromatic mole fractions compared to the assigned values (Figs. 2–3), indicating humidity passivation of active surface sites. Thus, losses were only apparent in their dry calibration standards (Fig. 4g). Compatible with this observation is the fact that the box plots (Fig. 3b and e) show a systematic underestima-

tion of aromatics only for NMHC_N_2, while for NMHC_air the results are more equally distributed.

Different hypotheses to explain losses of aromatics and C_6–C_8 alkanes did not result in simple and conclusive explanations. Losses were observed in individual systems when desorption was not sufficient, when adsorptive losses on inappropriate surfaces like newly installed stainless-steel lines (heated or not) occurred, or when dry sample gases were analysed. As long as a decrease in the C response is evident in both the calibration and NMHC_N_2, the submitted mole fractions did not differ much from the assigned values (e.g. YRK and AUC) (Figs. 4a, g, 2a, and i).

3.4.5 Chromatographic resolution

Poor peak separation or peak shape (tailing) influences the peak integration and the results. Both effects can mask other problems if the sample matrix is rather complex, such as in NMHC_air, where peak overlap is likely to occur in FID systems. Due to substantially different mole fractions in ambient air compared to NMHC_N_2, the chromatographic resolution, e.g. peak overlap, for NMHC_air differed considerably from the characteristics seen in NMHC_N_2.

Insufficient C_4–C_6 peak separation often resulted in mole fractions outside the 10 % class in NMHC_air, especially for 2-, 3-methylpentane; 2,2-,2,3-dimethylbutane; and 2-methyl-2-butene (Figs. 2 and 3b). Similar results were already reported in the AMOHA intercomparsion, where some participants had problems in separating 1-butene from 1,3-butadiene, *cis*-2-butene from 2-methylbutane, and isoprene from the methyl pentanes (Plass-Duelmer et al., 2006; Slemr et al., 2002). The reasons for the insufficient chromatographic separation include column degradation (AUC, FZJ_B), inadequate oven temperature programme (KOS), or non-baseline separation (HPB_A for C_5–C_6 alkanes) (for chromatograms see Supplement).

3.5 MS systems

Compared to FID systems, MS systems allow a better compound identification and peak separation at the cost of detector stability. With few exceptions, HPB_B (MS) reported the NMHCs within the 5 % class (Fig. 2q). It should be kept in mind that for HPB this was not a blind intercomparison. However, the ACTRIS mixtures were treated like unknown samples. Further, HPB_B was not used for the determination of the assigned values. The instrument is operated with a FID running in parallel to the MS detector. While the FID revealed stable behaviour of the instrument, in the MS signal drifts were observed by HPB. Thus, in routine measurements the MS is tuned weekly and every air sample is accompanied by a calibration measurement. In fact the HPB_B (MS) system was the best-performing MS system in this intercomparison, indicating that NMHC measurements within the 5 % class (ACTRIS DQOs) are achievable by MS systems.

The relatively large deviations from the assigned reference values in NMHC_N_2 and NMHC_air observed for CMN and the Medusa systems (Fig. 3) were mainly due to calibration issues (two-step calibration; see Sect. 3.3). Nevertheless, the very good repeatability of the Medusa systems indicates the potential to perform high-quality NMHC measurements within the 5 % class (Fig. 2u–w).

FZJ_A was optimised to perform fast chromatography as the instrument is employed in aircraft measurements. The sample volume is kept small in order to reduce the sampling time. With a chromatography time of 3 min, the peak resolution can hardly be compared to the other GC systems. Nevertheless, FZJ_A performed fairly well for normal alkanes and aromatics, whereas branched alkanes and alkenes showed larger deviations from the assigned values (Fig. 2p). Whether this was due to the rather complex 74-component calibration standard in the 0.1 to 10 nmol mol^{-1}-range (Apel-Riemer Environmental Inc.) cannot be judged from the available data. Furthermore, breakthrough of C_4 compounds was reported by FZJ_A. In general, the blank chromatogram revealed many peaks (chromatogram not shown), which possibly affected the results, especially in NMHC_air.

For NMHC_N_2 the MS systems of PUY and SIR reported most values with a deviation less than 10 %, whereas for NMHC_ air more of the reported values were outside the 10 % class (Fig. 2r and s). For PUY this was probably due to drifting calibration standard measurements (up to 20 %) and poor repeatability; for SIR it was probably connected to high blanks (relatively high blank values compared to assigned values (Table S7)) and poor stability of the calibration measurements.

The MS at SMR clearly underestimated the NMHC mole fractions in NMHC_N_2 (Fig. 2t), except for isoprene. In contrast, SMR reported all values within the 10 % class (Table S3) for NMHC_air. SMR reported a non-linear calibration curve and low reproducibility of the submitted calibration measurements, whereas the two NMHC mixtures were reproducibly measured.

In summary, the calibration, drift, and non-linearity are important issues for MS systems, which have to be handled with most care when using a GC-MS system for the measurements of NMHCs.

3.6 Other issues

During the ACTRIS intercomparison only very dry NMHC mixtures were analysed, and therefore a full performance assessment of water management systems (Nafion® Dryers, cold traps, or hydrophobic adsorbents at room temperature) cannot be made. Nevertheless, some basic conclusions can be drawn. The cold-trap systems used by YRK and HPB_A (Table S2a) exhibited no artefacts. Such systems sometimes have a large internal volume for water removal, and, whilst very suitable for online measurements, they are not so well suited for conditions where limited flushing vol-

ume is an issue, e.g. when analysing limited sample volumes. In this intercomparison, where dry samples were analysed, this method was superior compared to Nafion® Dryers where alkene artefacts are observed (see Sect. 3.4.3). The use of hydrophobic adsorbents at room temperature indicated no problems for HPB_B. However, the weak adsorbents used in HPB_B are not appropriate for the adsorption of low-boiling NMHCs (C_2–C_3).

Ozone management was not in the scope of this ACTRIS intercomparison study, and, furthermore, ozone is rapidly destroyed on metal surfaces; thus no ozone was present in the cylinders.

One specific issue was associated with the ZSF system, which had been brought to 2650 m a.s.l. shortly before this intercomparison. The reduced atmospheric pressure might have caused changes in the chromatographic conditions which had not been adjusted at the time of the measurements.

3.7 PTR-MS results

The two NMHC mixtures were analysed with the PTR-MS systems of SMR II and WCC-VOC. Isoprene in NMHC_N_2 fitted well inside the 5 % class, whereas isoprene in NMHC_air, toluene, and benzene in both NMHC mixtures were reported outside the 10 % class. Detailed results and some explanations are given in the Supplement.

3.8 Comparison with previous intercomparisons

During AMOHA phase 4 (Plass-Duelmer et al., 2006) and NOMHICE phase 4 (Apel et al., 2003) measurements of whole air and synthetic test samples were compared. As outlined in the Introduction, conditions were different and, accordingly, these studies cannot be compared with the ACTRIS intercomparison in the strictest sense. However, the whole air test samples supplied by canisters (NOMHICE and AMOHA phase 4 part 1) or sampled into individual canisters by participants (AMOHA phase 4 part 2) had a similar complexity to the whole air sample used in the actual intercomparison (e.g. 20–50 % of NMHCs < 0.1 nmol mol^{-1}). Originally introduced by Apel et al. (2003) and modified by Plass-Duelmer et al. (2006), a ranking procedure defining a score for quality and quantity of the provided results by each lab was introduced:

$$\text{Rank} = \frac{((n < +10\%) + 0.75\,(+10\% < n < +25\%) + 0.5\,(+25\% < n < 50\%))}{N} + 0.3\left(\frac{N}{X}\right) - 0.3(k),$$

where n is the number of reported values falling into the given reference intervals, N is the total number of compounds reported, X is the total number of compared compounds, and $k = \frac{\sum[\frac{x}{\text{ref}-1}]}{N}$ the averaged deviation of the reported values x from the reference values "ref". For reference values below 0.05 nmol mol^{-1} the bracket term is taken as $k = \frac{\sum \frac{x-\text{ref}}{50\text{pmol/mol}}}{N}$ (for details, see Plass-Duelmer et al., 2010). This "Rank" score can reach a maximum of 1.3 (all compounds measured and correct within 10 %) down to negative numbers for substantial deviations from the reference (large k). Minimum, median, and maximum ranks, respectively, for NOMHICE part 4 are 0.23, 0.81, and 1.16 (37 compounds); for AMOHA 4 phase 1 are 0.82, 1.02, and 1.14 and for phase 2 −0.31, 1.0, and 1.12; and in this study for NMHC_air are 0.49, 1.03, and 1.19 (the latter excludes results by the reference laboratories). The best-performing laboratories were in all studies similar at 1.14–1.19, the mid-quality increased from NOMHICE to AMOHA and this study, and the lowest-performing labs were best in AMOHA 4 phase 1 and ACTRIS. If we interpret the results as development over time, there is a tendency of improvement of the lower-performing labs, whereas the medium to best laboratories perform essentially unchanged over the last 15 years. However, AMOHA was a "learning" intercomparison with phases of increasing complexity and feedback to the participants in between, which in the end yielded the best performance for AMOHA 4 phase 1. Compared to this, ACTRIS may be seen as a snapshot with reasonable performance, as well as highlighting the need of more regular feed-back to the stations.

4 Conclusions

In the NMHC intercomparison exercise performed in the European infrastructure project ACTRIS, a significant number of instruments were capable of measuring NMHC in nitrogen (NMHC_N_2) fairly accurately: 88 % of the submitted NMHC values were within 10 %, and 58 % even within 5 %, of the reference values, which are the DQOs of ACTRIS with respect to the deviation to assigned values. It should be noted that NMHC_N_2 was almost identical to the NPL calibration standards used at the stations and a substantial number of deviations was not expected. Participants generally achieved very good repeatability in their measurements in line with the objectives of 2 %.

In compressed whole air (NMHC_air) generally more frequent and larger deviations to the assigned values compared to NMHC_ N_2 were observed (77 % of the reported values were within 10 %, but only 48 % were within 5 %). It should be noted that this comparison uses test gases which do only partly reflect the complexity of ambient air, e.g. no ozone and low water content. On the one hand, an important contributor to insufficient results in NMHC_air was blank issues observed in zero-gas measurements in some of the systems, especially those using a Nafion® Dryer. On the other hand, systems with cold traps exhibited smaller blank issues. The study highlights the importance of good zero-gas measurements to determine realistic blank values to be subtracted from measurement results.

Another factor contributing to the poorer NMHC_air results is the reduced chromatographic resolution, particularly in the range of C_4–C_6 compounds. Generally, those systems using direct calibrations in the nanomole-per-mole range achieved better results than those using whole air calibration standards. This confirms and emphasises the results found in the AMOHA and GAW intercomparisons (Plass-Duelmer et al., 2006; Rappenglueck et al., 2006; Slemr et al., 2002) as the two-step calibration and more complex matrix in whole air calibration standards introduce additional potential errors. For ethyne, losses may occur due to breakthrough in the adsorption trap, and yet unexplained reduced C response was observed in several systems. This intercomparison supports previous studies, finding that it is essential to calibrate ethyne directly and carefully characterise the response of the system in dry calibration standard and humid ambient air sample matrices. The use of FID C responses proved to be a powerful tool because it helped to identify problems in a number of analytical systems. However, as long as a system behaves similarly in different sample gas matrices, deviations in the C response may cancel, resulting in correct mole fractions. But this requires thorough testing of the respective GC systems. Breakthrough is generally an issue for C_2–C_3 hydrocarbons in adsorptive traps. Deviations from the expected C responses for low-boiling hydrocarbons were mainly observed in systems using the PerkinElmer Thermodesorber with air-toxics/air-monitoring traps. Whether these deviations were due to breakthrough or split injection issues could not be resolved. Almost all of the participating instruments indicated losses of C_7–C_8 aromatic compounds, most probably due to adsorptive losses. Despite such losses, many participants achieved good results for aromatics, but overall deviations were slightly larger than for other compound groups. On average, FID systems achieved better results, but good measurements were also obtained with GC-MS systems; however, since the MS is less stable than FID, more frequent calibrations are required.

Another important result of this intercomparison is that in more than 25 % of the reported results uncertainties were substantially underestimated and major uncertainty contributions were not correctly assessed. Last but not least, erroneous results were also caused by the occasionally inattentive data submission, with mistakes and incomplete information. While these problems were detected and resolved in the relatively small data set of this intercomparison, it is an issue with submission of insufficiently controlled data sets to public data centres and end-users.

The PerkinElmer Online Ozone Precursor Analyzer is the only commercially available instrument used by five participants in this intercomparison. Although these were not among the best performing in this study, reasonable results can be achieved. We demonstrated that the ACTRIS DQOs, albeit demanding, can be achieved with state-of-the art measurement systems. However, equally important for achieving high-quality results are experienced operators, comprehensive quality assurance and quality control, well-characterised systems, and sufficient manpower to operate the systems and evaluate the data.

The Supplement related to this article is available online at doi:10.5194/amt-8-2715-2015-supplement.

Acknowledgements. The research leading to these results was performed under the ACTRIS Research Infrastructure Project supported by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 262254. We acknowledge the Swiss National Air Pollution Monitoring Network (NABEL), the Swiss Federal Office for the Environment (FOEN), Empa, HALCIM (funded by FOEN), and the International Foundation High Altitude Research Stations Jungfraujoch and Gornergrat (HFSJG) for their support. The Department of Environment, Food and Rural Affairs (DEFRA) is acknowledged for the funding of the VOC measurements in the UK. The Umweltbundesamt Deutschland (Federal Environment Agency, Germany) is acknowledged for their support to the WCC-VOC. J. Jäger and R. Wegener would like to thank Deutsche Forschungsgemeinschaft for funding within the priority programme HALO (WE-4384/2-2). The LSCE team acknowledges the support from CEA, CNRSm, and IPSL. The Academy of Finland Centre of Excellence (grant no. 1118615) and the Nordic Centre of Excellence CRAICC are acknowledged for their support.

Edited by: N. Yassaa

References

Andreae, M. O. and Merlet, P.: Emission of trace gases and aerosols from biomass burning, Global Biogeochem. Cy., 15, 955–966, 2001.

Apel, E. C., Calvert, J. G., and Fehsenfeld, F. C.: The nonmethane hydrocarbon intercomparison experiment (NOMHICE) – Task-1 and Task-2, J. Geophys. Res.-Atmos., 99, 16651–16664, 1994.

Apel, E. C., Calvert, J. G., Gilpin, T. M., Fehsenfeld, F. C., Parrish, D. D., and Lonneman, W. A.: The Nonmethane Hydrocarbon Intercomparison Experiment (NOMHICE): Task 3, J. Geophys. Res.-Atmos., 104, 26069–26086, 1999.

Apel, E. C., Calvert, J. G., Gilpin, T. M., Fehsenfeld, F., and Lonneman, W. A.: Nonmethane Hydrocarbon Intercomparison Experiment (NOMHICE): Task 4, ambient air, J. Geophys. Res.-Atmos., 108, 4300, doi:10.1029/2002JD002936, 2003.

Badol, C., Borbon, A., Locoge, N., Léonardis, T., and Galloo, J. C.: An automated monitoring system for VOC ozone precursors in ambient air: Development, implementation and data analysis, Anal. Bioanal. Chem., 378, 1815–1827, 2004.

Barlow, R.: Statistics: a Guide to the Use of Statistical Methods in the Physical Sciences, Wiley, Chichester, 1989.

Bernardo-Bricker, A., Farmer, C., Milne, P., Riemer, D., Zika, R., and Stoneking, C.: Validation of speciated nonmethane hydrocarbon compound data collected during the 1992 atlanta inten-

sive as part of the Southern Oxidants Study (SOS), J. Air Waste Manage., 45, 591–603, 1995.

Bronštejn, I. N.: Handbook of Mathematics, Springer, Berlin, 2007.

Burns, W. F., Tingey, D. T., Evans, R. C., and Bates, E. H.: Problems with a Nafion membrane dryer for drying chromatographic samples, J. Chromatogr., 269, 1–9, 1983.

De Saeger, E. and Tsani-Bazaca, E.: EC Intercomparison of VOC measurements, in: Photo-Oxidants: Precursors and Products, edited by: Borell, P. M., Borell, P., Cvitas, T., and Seiler, W., SPB Academic Publishing bv, Den Haag, 1993.

Dietz, W. A.: Response factors for gas chromatographic analyses, J. Gas Chromatogr., 5, 68–71, 1967.

Faiola, C. L., Erickson, M. H., Fricaud, V. L., Jobson, B. T., and VanReken, T. M.: Quantification of biogenic volatile organic compounds with a flame ionization detector using the effective carbon number concept, Atmos. Meas. Tech., 5, 1911–1923, doi:10.5194/amt-5-1911-2012, 2012.

Gong, Q. and Demerjian, K. L.: Hydrocarbon losses on a regenerated Nafion Dryer, J. Air Waste Manage., 45, 490–493, 1995.

Grenfell, R. J. P., Milton, M. J. T., Harling, A. M., Vargha, G. M., Brookes, C., Quincey, P. G., and Woods, P. T.: Standard mixtures of ambient volatile organic compounds in synthetic and whole air with stable reference values, J. Geophys. Res., 115, D14302, doi:10.1029/2009JD012933, 2010.

Gros, V., Sciare, J., and Yu, T.: Air-quality measurements in megacities: Focus on gaseous organic and particulate pollutants and comparison between two contrasted cities, Paris and Beijing, CR Geosci., 339, 764–774, 2007.

Hahn, J.: Results of hydrocarbon interlaboratory comparison experiment in TOR, in: Transport and Transformation of Pollutants in the Troposphere, edited by: Borell, P. M., Borell, P., Cvitas, T., and Seiler, W., APB Academic Publishing bv, Den Haag, 399–403, 1994.

Helmig, D.: Ozone removal techniques in the sampling of atmospheric volatile organic trace gases, Atmos. Environ., 31, 3635–3651, 1997.

Helmig, D.: Air analysis by gas chromatography, J. Chromatogr. A, 843, 129–146, 1999.

Helmig, D. and Vierling, L.: Water adsorption capacity of the solid adsorbents Tenax TA, Tenax GR, Carbotrap, Carbotrap C, Carbosieve SIII, and Carboxen 569 and water management techniques for the atmospheric sampling of volatile organic trace gases, Anal. Chem., 67, 4380–4386, 1995.

Helmig, D., Tanner, D. M., Honrath, R. E., Owen, R. C., and Parrish, D. D.: Nonmethane hydrocarbons at Pico Mountain, Azores: 1. Oxidation chemistry in the North Atlantic region, J. Geophys. Res-Atmos., 113, D20S91, doi:10.1029/2007jd008930, 2008.

JCGM: Evaluation of measurement data – Guide to the expression of uncertainty in measurement, 134 pp., 2008.

Koppmann, R. (Ed.): Volatile Organic Compounds in the Atmosphere, Balckwell, Oxford, 2007.

Koppmann, R., Johnen, F. J., Khedim, A., Rudolph, J., Wedel, A., and Wiards, B.: The influence of ozone on light nonmethane hydrocarbons during cryogenic preconcentration, J. Geophys. Res., 100, 11383–11391, doi:10.1029/95JD00561, 1995.

Monks, P. S., Granier, C., Fuzzi, S., Stohl, A., Williams, M. L., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., Blake, N., Blake, R. S., Carslaw, K., Cooper, O. R., Dentener, F., Fowler, D., Fragkou, E., Frost, G. J., Generoso, S., Ginoux, P., Grewe, V., Guenther, A., Hansson, H. C., Henne, S., Hjorth, J., Hofzumahaus, A., Huntrieser, H., Isaksen, I. S. A., Jenkin, M. E., Kaiser, J., Kanakidou, M., Klimont, Z., Kulmala, M., Laj, P., Lawrence, M. G., Lee, J. D., Liousse, C., Maione, M., McFiggans, G., Metzger, A., Mieville, A., Moussiopoulos, N., Orlando, J. J., O'Dowd, C. D., Palmer, P. I., Parrish, D. D., Petzold, A., Platt, U., Pöschl, U., Prévôt, A. S. H., Reeves, C. E., Reimann, S., Rudich, Y., Sellegri, K., Steinbrecher, R., Simpson, D., ten Brink, H., Theloke, J., van der Werf, G. R., Vautard, R., Vestreng, V., Vlachokostas, C., and von Glasow, R.: Atmospheric composition change – global and regional air quality, Atmos. Environ., 43, 5268–5350, 2009.

Parrish, D. D. and Fehsenfeld, F. C.: Methods for gas-phase measurements of ozone, ozone precursors and aerosol precursors, Atmos. Environ., 34, 1921–1957, 2000.

Pérez Ballesta, P., Field, R., and De Saeger, E.: Interlaboratory exercises for volatile organic compound determination, Atmos. Environ., 35, 5729–5740, 2001.

Placet, M., Mann, C. O., Gilbert, R. O., and Niefer, M. J.: Emissions of ozone precursors from stationary sources: a critical review, Atmos. Environ., 34, 2183–2204, 2000.

Plass-Duelmer, C., Khedim, A., Koppmann, R., Johnen, F. J., Rudolph, J., and Kuosa, H.: Emissions of light nonmethane hydrocarbons from the Atlantic into the atmosphere, Global Biogeochem. Cy., 7, 211–228, 1993.

Plass-Duelmer, C., Michl, K., Ruf, R., and Berresheim, H.: C2–C8 Hydrocarbon measurement and quality control procedures at the Global Atmosphere Watch Observatory Hohenpeissenberg, J. Chromatogr. A, 953, 175–197, 2002.

Plass-Duelmer, C., Schmidbauer, N., Slemr, J., Slemr, F., and D'Souza, H.: European hydrocarbon intercomparison experiment AMOHA part 4: canister sampling of ambient air, J. Geophys. Res., 111, D04306, doi:10.1029/2005JD006351, 2006.

Rappenglueck, B., Apel, E., Bauerfeind, M., Bottenheim, J., Brickell, P., Cavolka, P., Cech, J., Gatti, L., Hakola, H., Honzak, J., Junek, R., Martin, D., Noone, C., Plass-Duelmer, C., Travers, D., and Wang, D.: The first VOC intercomparison exercise within the Global Atmosphere Watch (GAW), Atmos. Environ., 40, 7508–7527, 2006.

Read, K. A., Lee, J. D., Lewis, A. C., Moller, S. J., Mendes, L., and Carpenter, L. J.: Intra-annual cycles of NMVOC in the tropical marine boundary layer and their use for interpreting seasonal variability in CO, J. Geophys. Res. Atmos., 114, D21303, doi:10.1029/2009JD011879, 2009

Romero, R.: The first laboratory intercomparison of light hydrocarbons in EMEP, Institute of applied Environ. Res., NILU, Kjeller, Norway, 1–18, 1995.

Rudolph, J.: Measurements of nonmethane hydrocarbons in the atmosphere, in: Volatile Organic Compounds in the Troposphere, edited by: Koppmann, R. and Ehhalt, D. H., Schriften des Forschungszentrum Jülich, Jülich, 175–180, 1999.

Sawyer, R. F., Harley, R. A., Cadle, S. H., Norbeck, J. M., Slott, R., and Bravo, H. A.: Mobile sources critical review: 1998 NARSTO assessment, Atmos. Environ., 34, 2161–2181, 2000.

Scanlon, J. Z. and Willis, D. E.: Calculation of flame ionization detector relative response factors using the effective carbon number concept, J. Chromatogr., 23, 333–340, 1985.

Slemr, J., Slemr, F., Partridge, R., D'Souza, H., and Schmidbauer, N.: Accurate measurements of hydrocarbons in the at-

mosphere (AMOHA): three European intercomparisons, J. Geophys. Res., 107, ACH15.1–15.25, 2002.

Solberg, S.: VOC measurements 2010, Norwegian Institute for Air Research, Kjeller, Norway, 2012.

Solberg, S.: VOC measurements 2011, Norwegian Institute for Air Research, Kjeller, Norway, 2013.

Sternberg, J. C., Gallaway, W. S., and Jones, D. T. L.: Chapter XVIII: the mechanism of response of flame ionization detectors, in: Gas Chromatography: Third International Symposium Held under the Auspices of the Analysis Instrumentation Division of the Instrument Society of America, edited by: Brenner, N., Callen, J. E., and Weiss, M. D., Academic Press, New York USA, and London, UK, 231–267, 1962.

Volz-Thomas, A., Slemr, J., Konrad, S., Schmitz, T., Apel, E. C., and Mohnen, V. A.: Quality assurance of hydrocarbon measurements for the German Tropospheric Research Focus (TFS), J. Atmos. Chem., 42, 255–279, 2002.

Warneck, P. (Ed.): Chemistry of the Natural Atmosphere, Elsevier, New York, USA, 1988.

Westberg, H. and Zimmerman, P.: Analytical methods used to identify nonmethane organic compounds in ambient atmospheres, in: Measurement Challenges in Atmospheric Chemistry, edited by: Newman, L., American Chemical Society, Washington, DC, 275–290, 1993.

WMO: WMO Global Atmosphere Watch (GAW) Strategic Plan: 2008–2015, Geneva, Switzerland, 2007a.

WMO: A WMO/GAW Expert Workshop on Global Long-Term Measurements of Volatile Organic Compounds, Geneva, Switzerland, 2007b.