Atmospheric
Measurement
Techniques

An interlaboratory comparison to quantify oxidative potential measurement in aerosol particles: challenges and recommendations for harmonisation

Pamela A. Dominutti[1], **Jean-Luc Jaffrezo**[1], **Anouk Marsal**[1], **Takoua Mhadhbi**[1], **Rhabira Elazzouzi**[1], **Camille Rak**[1], **Fabrizia Cavalli**[2], **Jean-Philippe Putaud**[2], **Aikaterini Bougiatioti**[3], **Nikolaos Mihalopoulos**[3,4], **Despina Paraskevopoulou**[3], **Ian Mudway**[5,6], **Athanasios Nenes**[7,8], **Kaspar R. Daellenbach**[9], **Catherine Banach**[10], **Steven J. Campbell**[5], **Hana Cigánková**[11], **Daniele Contini**[12], **Greg Evans**[13], **Maria Georgopoulou**[7], **Manuella Ghanem**[14], **Drew A. Glencross**[5,6], **Maria Rachele Guascito**[12,15], **Hartmut Herrmann**[16], **Saima Iram**[17], **Maja Jovanović**[18], **Milena Jovašević-Stojanović**[18], **Markus Kalberer**[19], **Ingeborg M. Kooter**[20,†], **Suzanne E. Paulson**[10], **Anil Patel**[21,22,a], **Esperanza Perdrix**[23], **Maria Chiara Pietrogrande**[24], **Pavel Mikuška**[11], **Jean-Jacques Sauvain**[25], **Katerina Seitanidi**[7], **Pourya Shahpoury**[26], **Eduardo J. d. S. Souza**[16], **Sarah Steimer**[21,22], **Svetlana Stevanovic**[17], **Guillaume Suarez**[25], **P. S. Ganesh Subramanian**[27], **Battist Utinger**[19], **Marloes F. van Os**[20], **Vishal Verma**[27], **Xing Wang**[13], **Rodney J. Weber**[28], **Yuhan Yang**[28], **Xavier Querol**[29], **Gerard Hoek**[30], **Roy M. Harrison**[31,32], and **Gaëlle Uzu**[1]

[1]University Grenoble Alpes, CNRS, INRAE, IRD, INP-G, IGE (UMR 5001), 38000 Grenoble, France
[2]European Commission, Joint Research Centre (JRC), Ispra, Italy
[3]Institute for Environmental Research and Sustainable Development, National Observatory of Athens,
Lofos Koufou, P. Penteli, Athens, 15236, Greece
[4]Environmental Chemical Processes Laboratory, Department of Chemistry, University of Crete, Heraklion, 71003, Greece
[5]MRC Centre for Environment and Health, Environmental Research Group, Imperial College London,
86 Wood Lane, London, W12 0BZ, UK
[6]National Institute of Health Research, Health Protection Research Unit in Environmental Exposures and Health,
Imperial College London, London, UK
[7]Center for the Study of Air Quality and Climate Change, Institute of Chemical Engineering Sciences,
Foundation for Research and Technology Hellas, Patras, 26504, Greece
[8]Laboratory of Atmospheric Processes and their Impacts, Institute of Environmental Engineering,
École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland
[9]Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, 5232 Villigen PSI, Switzerland
[10]Department of Atmospheric and Oceanic Sciences, University of California at Los Angeles,
520 Portola Plaza, Los Angeles, CA 90095, United States
[11]Department of Environmental Analytical Chemistry, Institute of Analytical Chemistry, Czech Academy of Sciences,
Veveří 97, 60200 Brno, Czech Republic
[12]Institute of Atmospheric Sciences and Climate, ISAC-CNR, Str. Prv. Lecce-Monteroni km 1.2, 73100 Lecce, Italy
[13]Southern Ontario Centre for Atmospheric Aerosol Research, University of Toronto, Toronto, M5S 3E5, Canada
[14]Department of Pollutant Metrology, Institut National de Recherche et de Sécurité (INRS),
54500 Vandœuvre-lès-Nancy, France
[15]Department of Biological and Environmental Sciences and Technologies (DISTEBA),
University of Salento, 73100 Lecce, Italy
[16]Atmospheric Chemistry Department (ACD), Leibniz Institute for Tropospheric Research (TROPOS),
Permoserstraße 15, 04318 Leipzig, Germany
[17]School of Engineering, Deakin University, Geelong, 3216 Victoria, Australia
[18]Vinča Institute of Nuclear Sciences – National Institute of the Republic of Serbia,
University of Belgrade, 11 351 Belgrade, Serbia

[19]Department of Environmental Sciences, University of Basel, 4056 Basel, Switzerland

[20]TNO Environmental Modelling, Sensing and Analysis, Princetonlaan 6-8, 3584 CB Utrecht, the Netherlands

[21]Department of Environmental Science, Stockholm University, Stockholm, 11418, Sweden

[22]Bolin Centre for Climate Research, Stockholm, 11418, Sweden

[23]IMT Nord Europe, Institut Mines-Télécom, Univ. Lille, Centre for Energy and Environment, 59000 Lille, France

[24]Department of Chemical, Pharmaceutical and Agricultural Sciences, University of Ferrara,
Via Fossato di Mortara 17/19, 44121 Ferrara, Italy

[25]Center for Primary Care and Public Health (Unisanté), Department of Occupational and Environment Health (DSTE),
University of Lausanne, Lausanne, Switzerland

[26]Environmental and Life Sciences, Trent University, Peterborough, Canada

[27]Department of Civil and Environmental Engineering, University of Illinois at Urbana-Champaign,
205 North Mathews Avenue, Urbana, IL 61801, United States

[28]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, United States

[29]Institute of Environmental Assessment and Water Research (IDAEA-CSIC), 08034 Barcelona, Spain

[30]Institute for Risk Assessment Sciences, Utrecht University, Utrecht, 3584CM, the Netherlands

[31]Division of Environmental Health and Risk Management, School of Geography, Earth and Environmental Sciences,
Edgbaston, Birmingham, B15 2TT, UK

[32]Department of Environmental Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[a]now at: Bagchi School of Public Health, Ahmedabad University, Ahmedabad 380009, Gujarat, India

✝deceased

Correspondence: Gaëlle Uzu (gaelle.uzu@ird.fr)

Abstract. This paper presents the findings from a collaborative interlaboratory comparison exercise designed to assess oxidative potential (OP) measurements conducted by 20 laboratories worldwide. This study represents an innovative effort as the first exercise specifically aimed at harmonising this type of OP assay, setting a new benchmark in the field.

Over the last decade, there has been a noticeable increase in OP studies, with numerous research groups investigating the effects of exposure to air pollution particles through the evaluation of OP levels. However, the absence of standardised methods for OP measurements has resulted in variability in results across different groups, rendering meaningful comparisons challenging. To address this issue, this study engages in an international effort to compare OP measurements using a simplified method (with a dithiothreitol (DTT) assay).

Here, we quantify the OP in liquid samples to focus on the protocol measurement itself, while future international OP interlaboratory comparisons (ILCs) should aim to assess the whole chain process, including the sample extraction. We analyse the similarities and discrepancies observed in the results, identifying the critical parameters (such as the instrument used, the use of a simplified protocol, the delivery and analysis time) that could influence OP measurements and provide recommendations for future studies and interlaboratory comparisons even if other crucial aspects, such as sampling PM methods, sample storage, extraction methods and conditions, and the evaluation of other OP assays, still need to be standardised. This collaborative approach enhances the robustness of the OP DTT assay and paves the way for future studies to build on a unified framework. This pioneering work concludes that interlaboratory comparisons provide essential insights into the OP metric and are crucial to move toward the harmonisation of OP measurements.

1 Introduction

Over the last decade, many studies demonstrated associations between exposure to ambient air pollution and adverse human health outcomes (Hart et al., 2015; Laden et al., 2006; Lepeule et al., 2012; WHO, 2013, 2021). Adverse health effects attributable to particle matter (PM) are complex and diverse. Among environmental factors, PM is considered to be the largest contributor to morbidity and mortality globally (WHO, 2024). The causal mechanisms underpinning these adverse associations are diverse, with oxidative stress and inflammation (Leni et al., 2020; Li et al., 2003, 2008), genomic alterations (Huang et al., 2022), and damage to the nervous system function (Wilker et al., 2023), as well as epigenetic alterations and cognitive decline (Nicholson et al., 2022; Zare Sakhvidi et al., 2022), all cited as potential contributing pathways. Across these broad domains, the capacity of particles

and particle-derived chemicals to cause damaging biological oxidations appears to be the ultimately response of oxidative stress, both through the introduction of pro-oxidants and stable free radicals into the body, but also through secondary radical/oxidant generation through altered metabolism and induction of inflammation (Li et al., 2008). By definition, oxidative stress is a condition where excess production of reactive oxygen species (ROSs) and nitrogen species (RNSs) overwhelm endogenous antioxidant defences (Shankar and Mehendale, 2014). Generally, ROS and RNS production in the cells is regulated within physiological limits through the actions of antioxidant enzymes (e.g. superoxide dismutase, catalase) and low-molecular-weight water-soluble (e.g. glutathione, ascorbate) and fat-soluble (vitamin E) antioxidants (Alkoussa et al., 2020). This antioxidant system plays a valuable key role by limiting ROS and RNS damage, which is associated with cytotoxicity and the induction of inflammation due to changes in the cellular redox balance (Cassee et al., 2013; Gao et al., 2020; Kelly and Fussell, 2017; Sies, 2018). The capacity of PM to invoke biological oxidations has therefore been proposed as a proxy measure of their toxicity and has been referred to as their oxidative potential (OP), which is intrinsic to their possession of pro-oxidants or encompasses their capacity more holistically to simulate ROSs and RNSs through interaction with cells (Ayres et al., 2008; Bates et al., 2015; Cho et al., 2005; Sauvain et al., 2008; Uzu et al., 2011).

Consequently, the OP of PM is increasingly being studied as a potentially health-relevant metric to evaluate effects due to exposure to PM, in addition to PM mass concentration, in multiple regions across the globe (Bates et al., 2019; Bhattu et al., 2024; Daellenbach et al., 2017; Weichenthal et al., 2019). OP is a relatively simple estimation of PM redox activity that reflects a complex interplay of all physico-chemical properties (chemical composition, surface-area, solubility and particle size) contributing to the ROS and RNS generation and the oxidation of target biomolecules or probes. Implicit within this approach is the contention that not all constituents of ambient PM are equally as harmful and that those that drive damaging redox reactions, either directly or indirectly, present a greater hazard. Thus, PM composition should be considered a factor more directly linked to adverse health effects than PM mass concentration, highlighting the need to study additional health-relevant metrics such as the OP (Park et al., 2018).

In last decades, there has been an increased interest in measuring and developing OP studies, applying different in vivo or in vitro assays and aerosol characterisation techniques to estimate the main sources of OP related to PM (Guascito et al., 2023), and attempting to integrate this metric into health studies. Several acellular chemical methods have been applied for the estimation of the OP of atmospheric particles since these assays allow faster measurement and are less labour-intensive than cell culture or in vivo methods (Bates et al., 2019). In addition, these assays aim to mimic the interaction between PM and different lung antioxidants (e.g. glutathione, ascorbate) during inhalation. The acellular assays which are most commonly applied include several probe approaches based on antioxidants or surrogates, such as the dithiothreitol (DTT) assay, ascorbic acid (AA) assay, glutathione (GSH) assay, ferric–xylenol orange (FOX) assay, 9,10-bis(phenylethynyl)anthracene-nitroxide (BPEAnit) ROS assay and 2,7-dichlorofluorescein (DCFH) assay for bound particles. Molecular probes display variable sensitivities to PM components due to their unique redox potentials and chemical reaction routes, contributing to aerosol OP values. Therefore, it may be necessary to use several assays simultaneously for a broader assessment of the chemical species in PM potentially triggering oxidative stress and to evaluate which of these probes might be most indicative and closely linked to health effects. Furthermore, one of the main challenges within this rapidly expanding research field is the diversity of analytical methods and protocols used for each OP assay, which require standardised protocols to support synthesis across the evidence base (Ayres et al., 2008).

In 2008, a previous workshop gathered experts on OP and developed consensus statements addressing the importance of standardised samples, the comparison between oxidative potential tests, the formulation of consistent standard test protocols, and the establishment of connections between OP tests and epidemiological findings as reliable predictors of adverse health outcome (Ayres et al., 2008). Despite more than 15 years having elapsed since that workshop, whilst protocols have matured, evolved and proliferated worldwide, little concrete work has been performed regarding the harmonisation and standardisation of these methods.

One of the main objectives of the RI-URBANS European project (https://riurbans.eu/, last access: 30 December 2024) is to bring accessible service tools to enhance air quality monitoring networks, including evaluating air pollution exposure. As OP has been proposed and recommended as a parameter to be measured in the proposal for a new European Air Quality Directive (Council of the European Union, 2024), an international OP interlaboratory comparison (ILC) was launched to assess the consistency of OP measurements between participants that apply different OP DTT protocols, hindering comparison of results obtained worldwide. The main goal of the ILC was to identify potential discrepancies in results (obtained with the OP DTT assay, one of the most common acellular assays used for measuring the OP) that may arise due to differences in experimental procedures, equipment or analytical techniques. This ILC constitutes a first step to identify potential sources of variability, resulting in the enhancement of the overall accuracy, reliability and comparability of OP measurements.

This paper presents the set-up and results of this first large ILC study based on the dithiothreitol (DTT) assay, with a large number of participants (20 groups). The first section includes a description of how a simplified protocol was obtained, along with the coordination and management of the

ILC. Subsequently, the results are presented in the second section, combined with statistical analysis, for both the harmonised protocol and the home protocols of each participant. Finally, some major findings and recommendations are presented.

2   Intercomparison strategy

This ILC was proposed within the RI-URBANS European project framework to evaluate the discrepancies and commonalities of OP measurements obtained by the different participating laboratories. The set-up of the protocol was led by a working group of laboratories with considerable experience in oxidative potential: FORTH (the Foundation for Research and Technology – Hellas (FORTH, Greece); NOA (National Observatory of Athens, Greece); ICL (Imperial College London, United Kingdom); IGE (Institute of Environmental Geosciences, France); and UoB (University of Birmingham, United Kingdom) (i.e. the "core group"). Considerations regarding sampling techniques of PM filters or monitoring strategic approaches are beyond the scope of this exercise.

Multiple OP assays are available; however, following a literature review, it was decided to prioritise the DTT assay for this first ILC due to its widespread adoption and long-term application facilitating broad participation from laboratories. The core group first produced a harmonised and simplified method, detailed in a standardised operation procedure (SOP), that was integrated, implemented and tested by IGE, the organiser for this ILC. This SOP is called the "RI-URBANS DTT SOP" in the following and is presented in detail in Sect. 2.2. Section 2.1 presents the selection procedure for some parameters of the DTT assay according to variations observed in the literature. Section 2.3 to 2.7 comprise different parts of the implementation of the ILC, along with the procedure for data processing.

2.1   Testing the parameters for implementing the simplified RI-URBANS protocol

The simplified RI-URBANS DTT SOP was adapted from the original DTT protocols published in the early 2000s (SOP1: Li et al., 2003, 2009; SOP2: Cho et al., 2005; SOP3: Kumagai et al., 2002; called SOP1, SOP2 and SOP3, hereafter). The principle of the DTT assay relies on the production of superoxide radicals, with DTT acting as a surrogate for cellular reducing agents. This probe contains thiol groups similar to GSH and is subject to oxidation, forming stable cyclic disulfides by donating electrons to oxygen through intermediate redox-active species from PM. In the assay proposed by Kumagai et al. (2002), PM is incubated with DTT and its rate of oxidation over time is assessed through the use of 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) with the reaction product 2-nitro-5-thiobenzoic acid (TNB), being detected at

412 nm. Whilst DTT is in excess, the rate of DTT oxidation is proportional to the concentration of redox-active species in the PM (Cho et al., 2005; Sauvain et al., 2008). The DTT loss over time can then be expressed per unit concentration of PM (usually µg) or by volume of air (m^3) to provide a measure of the intrinsic reactivity of the particles and assess human exposure (Cho et al., 2005), respectively.

A review of the pioneer DTT assay protocols revealed differences, and therefore to derive a simplified protocol, some variations in the parameters were examined by the ILC organiser laboratory. These results were evaluated by the RI-URBANS core group to obtain the final harmonised protocol.

Three key parameters were tested. First, the necessity for the inclusion of trichloroacetic acid (TCA) was evaluated. In the original DTT protocol by Kumagai et al. (2002), the reaction between samples and DTT is quenched at a specified incubation time by the addition of 1.0 mL of 10 % TCA to the incubation mixture. Subsequently, 0.5 mL of the reaction mixture is extracted and combined with DTNB and Tris-HCl buffer (pH 8.9). However, more recently, Li et al. (2009) demonstrated that DTNB rapidly reacts with DTT, with the absorption reaching its maximum value immediately and remaining stable for over 2 h. In our initial tests, we found that this parameter was consistent, showing no differences over all the samples tested with or without TCA (Fig. S1 in the Supplement). Thus, in the simplified OP RI-URBANS SOP, TCA addition was omitted, and we introduced the DTNB solution directly into the mixture at the prescribed incubation times and recorded the absorption after the 30 min reaction.

Second, ethylenediaminetetraacetic acid (EDTA) is present in some OP DTT protocols, whether in the buffer of the incubating PM sample with DTT (Kumagai et al., 2002; Li et al., 2009) or in the titration mixture of DTNB (Cho et al., 2005; Kumagai et al., 2002). EDTA is a strong chelating agent which is widely used in biological assays to prevent microbial contamination and to facilitate cell lysis and the extraction of cellular components. Using EDTA is helpful in purifying buffers at a low cost and decreasing a high rate of DTT loss in the blank by scavenging metal ions. However, it can lead to artefacts during the assay. This is especially critical when used in the reaction mixture with the PM, where it could induce complexation with redox changes. Moreover, this is particularly relevant for iron where complexation increases solubility and the resultant EDTA–Fe complex is redox-active (Gao et al., 2024) or, in the opposite way, EDTA can chelate some metallic species, preventing their reactivity (Charrier and Anastasio, 2012). In addition, EDTA also has antioxidant properties, which may compete in the solution with DTT (Thbayh et al., 2023). Our results for different samples show that the presence of EDTA in the buffer leads to underestimation of the OP DTT loss rates in comparison to that without EDTA. An impact was observed in the solutions, mainly the copper reference solution (1 µM) and an ambient PM filter for testing both

protocols (Fig. S2). That augmentation could also be related to an increase in the blank absorbance without EDTA due to the impurities in the buffer, but this is controlled by subtracting the blank. Little or no impact was shown in the 1,4-naphthoquinone solutions, an organic component. To prevent such undesirable interactions, many laboratories have introduced Chelex® 100, a sodium-form resin to purify the buffers used in the OP DTT assay (Calas et al., 2017, 2018; Charrier et al., 2015; Charrier and Anastasio, 2012), or used commercially available high-purity (e.g. LC-MS/MS grade) water and buffer mixture to create the OP reaction medium (Shahpoury et al., 2019, 2022). Nevertheless, the Chelex® commercial resin comes with a basic pH and requires a pre-treatment that could be an extra source of error for this current, first international intercomparison. Finally, EDTA was not included in the SOP to prevent complex ligand chemistries and because the use of Chelex® was too complex to be introduced for the first such ILC; high-grade buffer powders were sent to all the participants instead.

Third, original protocols include the use of a Tris-HCl buffer at pH 8.9 in the solution of titration with DTNB, and this remains widely used. However, Li et al. (2009) showed that the pH of the solution drives both the catalytic redox reaction rate and also the molar extinction coefficient of the product of the reaction, TNB. This report confirmed the previously obtained results by Danehy et al. (1971) that showed that DTNB suffers from alkaline decomposition above pH 8, increasing absorbance values (Fig. S3). As a result, for the simplified RI-URBANS DTT SOP, the core group selected a potassium phosphate buffer at a physiological pH of 7.4 to replace Tris-HCl in the DTNB solution. This prevents the DTNB alkaline decomposition because the pH is in the favourable range of 5.5–8 where TNB is in the TNB^{2-} form and the DTT+DTNB system does not show significant pH-dependent changes in the absorbance values (Li et al., 2009).

Finally, the simplified RI-URBANS DTT SOP also includes a variation considering the instrument used for the measurements. The SOP was elaborated and tested for both cuvette and plate-reader spectrophotometers, including the potential application of automatic samplers. However, the simplified SOP remains to be tested for other instruments (such as an liquid waveguide capillary cell (LWCC) instrument).

2.2 Simplified RI-URBANS DTT SOP and other measurements

The simplified RI-URBANS DTT SOP is proposed in Sect. S1 in the Supplement. It contains a first step for the preparation of the reagents needed for the analyses included in the ILC. A second step describes how to perform a calibration of the analytical device, using a DTT calibration curve with at least four points for a concentration range between 0 and 60 µM (titration with 1 mM DTNB and reading of TNB

formation at 412 nm). In the third step, the SOP defines the performance of the measurements for the test samples provided, including the assessment in triplicates of each test sample and control point (blanks). It should be noted that the DTT protocol also integrates the analysis of control points (blanks), which allows quantifying the inherent DTT background oxidation.

The duration of the analyses required for the ILC is variable depending on the instrument used. About 30 min completion time is required, including the assessment of ILC test samples and control points, when performed with a plate reader, and a similar time is needed to perform the calibration curve. When a cuvette-type spectrometer is used, the total analysis time can be at least 2 h to perform all the triplicates and the corresponding calibration curve.

Furthermore, apart from the analytical equipment for monitoring the chemical reactions, the simplified protocol also requires some standard laboratory equipment and conditions, such as access to ultra-pure water (18.2 MΩ cm^{-1}, total organic carbon (TOC) < 5 ppb) for the preparation of reagents, the use of a vortex for the homogenisation of samples, refrigerated baths to conserve the reagents, and transparent 96-well plates and dark tubes. The samples also need to be kept under agitation during the experiment and at a constant temperature of 37.4 °C.

The calculation of the chemical reaction rate of OP DTT during this ILC involves a conversion using a calibration curve. Once the results are obtained, the kinetics of the DTT oxidation can be calculated by subtracting both the intrinsic absorption of each sample (absorption obtained from the samples before the addition of reagents to remove a potential matrix effect between samples) and the inherent DTT auto-oxidation rate (slope of a control sample) from the DTT consumption rate in the presence of PM.

Participants were asked to perform additional OP measurements on the same samples, following the protocol in use in their laboratory ("home protocol(s)" or DTT-home) if they wanted to do so. In this case, they had to provide the results in the particular units requested, depending on the applied assay.

2.3 Type of samples – test materials

The ILC was performed using three samples (SP1–SP3), including different concentrations of ambient PM and a commercial positive control (1,4-naphthoquinone; CAS 130-15-4, Sigma Aldrich), all extracted in ultra-pure water and prepared by the ILC organiser. Providing ultra-pure water extracts – instead of filter fragments – to the participants was selected for the current first ILC to avoid additional uncertainties associated with the use of different procedures of sample extraction, different quality of the ultra-pure water for sample extractions and changes linked to the processing equipment available. More specifically, the samples sent to the participants included SP1 – 1,4-naphthoquinone so-

lution (reference compound, at $5\,\mu g\,mL^{-1}$); SP2 – extract from a PM sample influenced by biomass burning emissions (obtained from a chamber experiment, at $25\,\mu g\,mL^{-1}$); and SP3 – an urban PM extract highly influenced by traffic emissions (from pooled roadside samples obtained from TEOM–FDMS (tapered element oscillating microbalance – filter dynamics measurement system) reference samplers, at $25\,\mu g\,mL^{-1}$). Additionally, a fourth sample, SP4 – a sample extracted from a blank/clean quartz filter, was sent to the participants, but it was not included in the evaluation since the measured values were close to the instrument limits of detection for most participants. For each of the four samples, all the sub-samples, distributed to participants, resulted from a unique 1 L solution obtained from the original sample substrate. For instance, SP1 and SP3 were powder and concentrated extraction, respectively, that were solubilised and homogenised for 75 min by vortex agitation in ultra-pure water. SP2 and SP4 were quartz fibre filters subjected to a 75 min vortex extraction in ultra-pure water.

Several 5 mL sample aliquots in dark polypropylene tubes were sent to each participant, according to their needs, allowing them to perform triplicate measurements for the RI-URBANS DTT SOP and all the home OP protocols implemented in their labs.

Solid potassium dihydrogen phosphate (CAS 7778-77-0, Roth), di-potassium hydrogen phosphate (CAS 7758-11-4, Roth), 5,5'-dithiobis(2-nitrobenzoic acid) (CAS 69-78-3, Roth) and 1,4-dithiothreitol (CAS 3483-12-3, Roth) were also distributed to the participants to prepare the solutions for the RI-URBANS DTT SOP, including the DTT solution, the dithiobis(2-nitrobenzoic acid) (DTNB) solution and the potassium phosphate buffer solution.

2.4 Transport of samples, ILC performance and duration

Test samples (SP1–SP4) were shipped to all participants on 17 January 2023 via courier in refrigerated and isolated ice packs and received as chilled liquid samples. The parcels were delivered between 18 January and 2 February 2023. On average, the participants received the samples in $3 \pm 2.5\,d$ and performed the analysis in $14 \pm 10.5\,d$ and up to 31 d after the reception of samples. The recording of these parameters allows their integration into the multiple linear models used in this work.

2.5 Reporting of the results

Participants were asked to report OP DTT results from the RI-URBANS SOP in nanomoles DTT per minute per microgram ($nmol\,DTT\,min^{-1}\,\mu g^{-1}$) and the percent of DTT consumed in micrograms per minute ($\mu g^{-1}\,min^{-1}$), applying three decimal digits for all three replicates of test samples. An Excel spreadsheet with all the calculations pre-included was prepared by the ILC organiser and shared with all the

participants to avoid calculation errors and to facilitate the standardisation of results. In addition, participants were invited to report, in the same format, the values for other OP tests, such as OP DTT-home, and other OP tests like AA, DCFH, OH, ESPR (electron spin (or paramagnetic) resonance), GSH and RP (redox potential) (routinely applied by each participant) on the same samples.

2.6 Number of participants

It is worth noting that for the first time, a total of 20 research groups participated in the exercise: 14 of them from Europe, including the United Kingdom, Italy, France, Switzerland, Greece, Germany, Serbia, the Czech Republic, the Netherlands and Sweden; 3 participants from the United States; 2 participants from Canada; and 1 from Australia. The participants were invited for their contribution to the RI-URBANS project through a public call to participate or because they contacted the ILC organiser directly and were selected due to their active role in the OP scientific community. The ILC was performed using anonymous participation; thus, a number was randomly assigned to each participant to present the results. Participant L5 cancelled their participation in the ILC, and two participants (L3 and L16) did not send their results for RI-URBANS DTT SOP.

2.7 Data evaluation

The ILC results were analysed by the European Commission Joint Research Centre (JRC), which provides independent, evidence-based science and knowledge support for EU policies and in conducting ILC exercises. The participation of an external independent evaluation was required following the International Global Standard ISO 5725-2, related to the accuracy of measurement methods and results ("Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method"). The data evaluation includes the assessment of the test sample homogeneity and stability, of each laboratory's repeatability, and of participants' performance.

2.7.1 Estimation of the assigned value and participant performance

Participant performance is evaluated based on biases against pre-assigned criteria including the assigned values and standard deviations for the proficiency test.

The choice of the methods for determining these assigned values and standard deviations is under the responsibility of the ILC organiser. The standard uncertainty of the assigned value should be as small as possible to minimise the risk that participants will receive underperformance signals because of inaccuracy in the determination of the assigned value. Three methods including the use of (i) a simple mean and standard deviation including all participants, (ii) the Q/Ham-

pel test, and (iii) the consensus value and standard deviation from expert laboratories were compared.

Six participants (L2, L4, L8, L12, L13 and L19) were selected a priori (without knowing their results) as expert laboratories based on their previous pioneering experience in developing OP measurement protocols and their strong expertise (i.e. > 10 publications) in the field. The "consensus among expert laboratories" approach was selected for determining the assigned values ($X = 0.53$, 0.14, 0.07 nmol min^{-1} µg^{-1} for SP1, SP2 and SP3, respectively) and standard deviations ($\sigma^* = 0.16$, 0.06, 0.04 nmol min^{-1} µg^{-1} for SP1, SP2 and SP3, respectively) for the proficiency test because it led to the smallest relative uncertainties for the assigned values, ranging from 10 % to 13 % for test samples SP1 to SP3.

The performance of each participant was further evaluated using z scores, a metric indicating the deviation of each data point from the assigned value as compared with the standard deviation for proficiency assessment, $z = (x_i - X)/\sigma^*$, where x_i is the result from participant i. An "action signal" is triggered if a participant's entry produces a z score exceeding $+3z$ or falling below $-3z$, indicating a deviation of more than 3 standard deviations from the assigned value. Similarly, a "warning signal" is raised for a participant z score above $+2z$ or below $-2z$, representing a deviation between 2 and 3 standard deviations. A participant z score between $-2z$ and $+2z$ signifies satisfactory performance concerning the standard deviation for proficiency assessment.

2.7.2 Analysis of the variability in the results

Additionally, the statistical distribution of results was evaluated using multiple linear regression models. In the first model, the effects of the protocol variables on the measured OP were investigated using linear regression models adjusted on the instrument (three class variables: plate reader, cuvette and LWCC), delivery time (time between sample shipment and reception, continuous) and analysis time (time between reception and analysis, continuous). An additional model (M2) compared the RI-URBANS SOP and the DTT-home protocols was further adjusted based on the protocol (two class variable: RI-URBANS, DTT-home). Finally, the evaluation of the average performances (three class variables: low $- 0 < | z$ score $| < 2$; middle $- 2 < | z$ score $| < 3$; and high $- | z$ score $| > 3$) was added in the M2 model to assess whether performance affected DTT activity in the same direction (i.e. positive or negative) while considering other protocol variables. Each model was run separately for SP1, SP2 and SP3 samples. All analyses were conducted using R (version 4.2).

3 Results and discussion

Out of the group with a total of 20 participants, 18 presented results obtained using the RI-URBANS DTT SOP. Different instruments were used to apply the simplified SOP. Overall, nine participants (47.5 %) used the cuvette-type spectrophotometer, eight (42 %) used the plate-reader-type spectrometer, and two (10.5 %) implemented LWCC measurements (one participant used two instruments).

3.1 Homogeneity of the samples

An initial assessment of the homogeneity of the OP measurements with the three test samples was performed by the ILC organiser using a plate-reader-type protocol, deriving the mean and the standard deviation of 10 replicate analyses performed on the same day. The results obtained from the coefficient of variations showed the sample variabilities were up to 12 %, 7 % and 9 % for SP1, SP2 and SP3, respectively (Fig. S5), showing a higher variability for the 1,4-naphthoquinone solution (SP1) compared to the two filter extracts (SP2 and SP3).

The overall uncertainty of the OP DTT assay has been evaluated to 18 % for PM$_{10}$ and 16.3 % for PM$_{2.5}$ by Molina et al. (2020). Despite some differences observed between our samples, the results are deemed acceptable, presenting a variability of around 10 %, which indicates a very good performance of the analysis.

3.2 Ageing of samples

To reduce the number of parameters affecting the preparation of the sample solutions, liquid solutions of each sample were prepared and sent to the different participants. However, liquid samples can undergo ageing processes, impacting OP levels over time. For this purpose, an ageing test was performed by the ILC organiser to evaluate the potential changes over time. It consisted of regularly implementing the RI-URBANS DTT SOP to obtain the values of each test sample over time. Figure S6 shows these results, where SP2 and SP3 do not show a strong change over time, while sample SP1 presents a pronounced ageing effect. In routine tests of the ILC organiser, the 1,4-naphthoquinone mother solution at such a high concentration is usually stable in a glass container for weeks, but here, a potential interaction with the PP tubes' inner surface may have happened. Consequently, ageing could be a variable of importance for the participants who analysed the samples toward the end of the required period, and such a parameter (date of analysis) was thus included in the parameters to be tested for the research of critical parameters.

3.3 Statistical distribution of results: participants' variability

In order to assess the intra-laboratory variability in the results, the coefficient of variation (COV = standard deviation / mean $\times$ 100) values of the results for each sample and each laboratory are presented in Fig. 1, while the standard deviations of the replicates reported for each sample are pre-
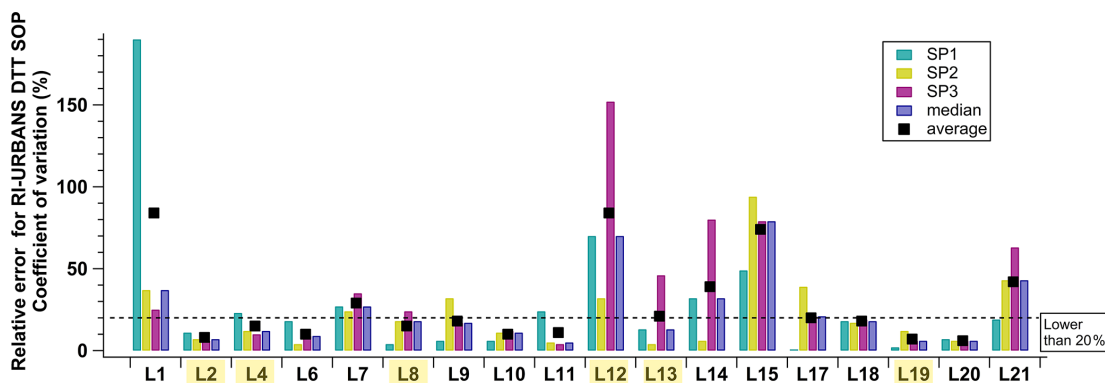
Figure 1. Coefficients of variation of each participant (L1 to L21) for the three samples in triplicates tested using the RI-URBANS DTT protocol and the median and mean repeatability for each laboratory. Yellow-highlighted participants are the ones selected for the calculation of the assigned values. The dashed line indicates the participants with a COV lower than 20 %.

sented in Fig. S7 and Table S1. Overall, higher COVs are observed for SP3 and SP1, where most participants (44.4 % and 38.9 %, respectively) presented higher values for this sample compared to the SP2 sample. Specifically, high average COV values are observed by L1, L12, L14, L15 and L21, with an average variation higher than 40 %. Only a few participants (six groups) presented a variation lower than 10 % for the three samples. This is the same pattern for the results obtained by the ILC organiser during the homogeneity test (see Sect. 3.1), where the COV for SP1 was larger than those for SP2 and SP3 but with the highest COV below 15 %. These findings confirm more homogenous results for samples SP2 and SP3 compared to SP1 but could also indicate that some participants failed to achieve repeatability observed by the ILC organiser. However, some groups (L2, L10, L19, L20) were able to produce very homogeneous results with COV < 10 % for all three samples.

3.4 Laboratory performances

The assessment of laboratory performances first presents the bias in results across participant groups compared to the assigned values and their associated standard deviation for each sample. As illustrated in Fig. 2, SP1 exhibited the highest variances, ranging from 130 % to −35 %, with only five groups displaying differences within ±10 %. The distribution of results for SP1 indicated a mix of overestimations and underestimations. For SP2, differences are within a narrower range from 43 % to −7 %, primarily favouring overestimations. For this sample, 12 participants returned results that were within ±10 % of the assigned value (see highlighted laboratory numbers in Fig. 2). Finally, the results for SP3 demonstrated the least variation among participants, with differences ranging from 30 % to −6 % and 16 participants within ±10 % of the assigned value, again favouring overestimation compared to the assigned value. In total, 14 laboratories obtained data with a ±10 % difference from the

assigned value for SP2 and SP3 (see highlighted laboratory numbers in Fig. 2). These results show again that the reference samples with 1,4-naphthoquinone (SP1) most probably present some characteristics leading to this variability and may be associated with a less-stable solution or led to the saturation of some detectors considering the relatively high concentrations, while the samples from filter extractions do not. Additionally, it is interesting to note that there is apparently no systematic pattern where a given participant would obtain out-of-range results for all samples. The results are really diverse, and most participants can obtain "acceptable" results for one or two samples (SP1–SP3) and larger variability associated with one "unacceptable result" for one of them. In 2020, Molina et al. (2020) explored the total uncertainty in OP DTT of a collection of samples and evaluated it to be 18 % for PM_{10} and 16.3 % for $PM_{2.5}$. The leading factors identified were the DTT consumption rate (regression and repeatability of experimental data) and the extraction volume operations (pipette). This underscores the need for further investigations on the experimental causes of the variations observed, possibly in the next ILC.

The individual performance of each group was further evaluated using z scores. The results are presented in Fig. 3. All underestimations fall within the acceptable range (lower than −2z). Additionally, it is noteworthy that no laboratory exhibits unsatisfactory performances across all three samples; for almost all participants, while one sample can present poor results, it coexists with two acceptable ones. This has strong implications for spatial and temporal analyses often conducted for OP, emphasising the need to assert that there are no systematic biases in the analyses. While factors like sample inhomogeneity may be playing a role (particularly for SP1), some other issues, including variability in the performance of the analysis, may have an impact. Hence, participants exhibiting significant deviations ($|z$ scores $| > 2$) for some of their results should thoroughly examine their proce-
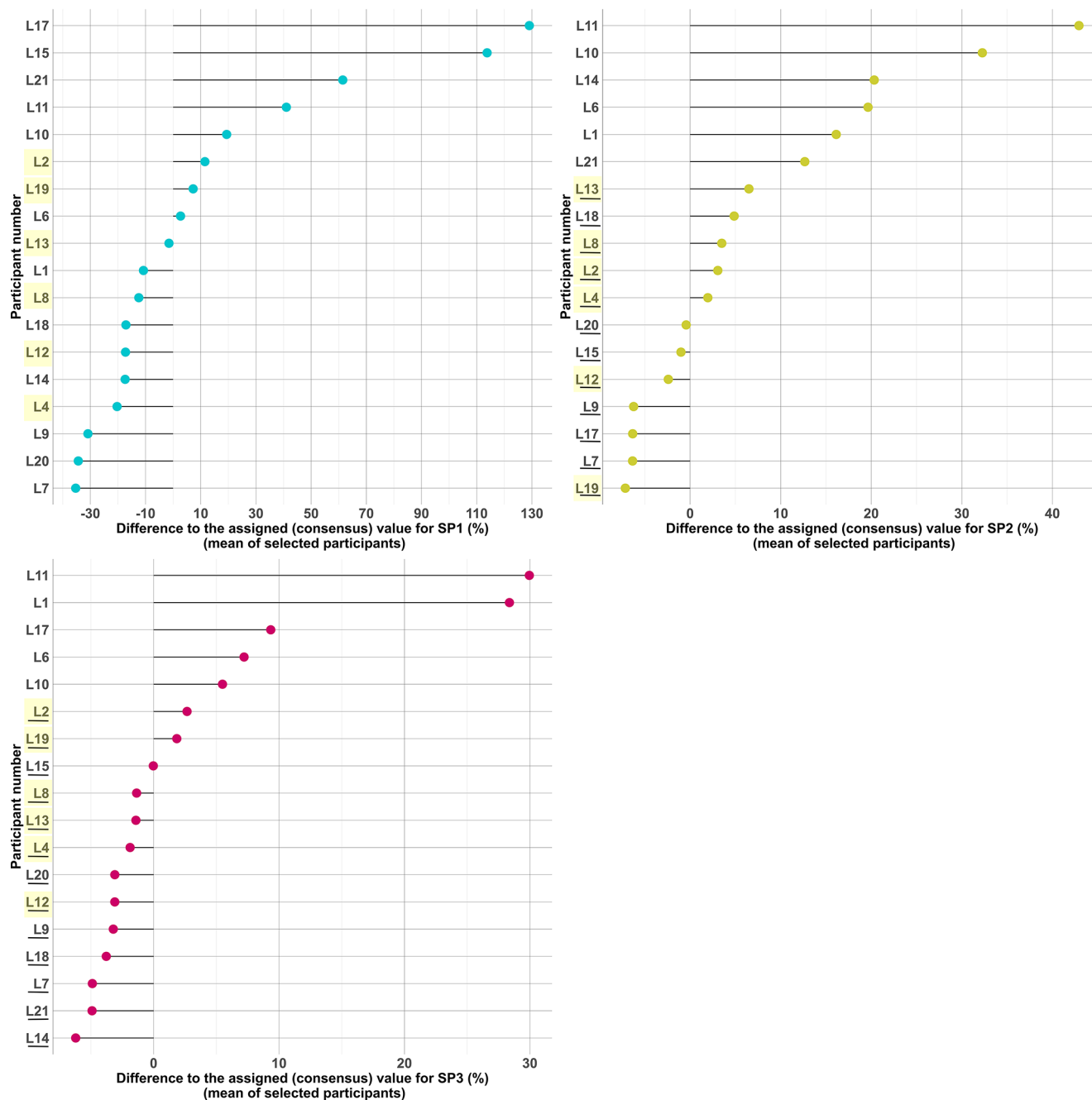
Figure 2. Percentage differences from the assigned (consensus) value for each sample (SP1, SP2 and SP3). The results compared the average of the triplicates reported by the participants. Yellow-highlighted participants are the ones selected for the calculation of the assigned values and the underlined ones are those that obtained data with a $\pm 10\%$ difference from the assigned value for SP2 and SP3.

dures and possibly implement appropriate corrective actions to avoid similar outcomes in future ILCs.

However, half of the participants achieved results within the acceptable limits of this test. Despite disparities, these findings are really promising, especially considering that this is the first intercomparison of its kind. For instance, such results are in the same range as those obtained for some of the first ILCs for polycyclic aromatic hydrocarbons (PAHs; Grandesso et al., 2012; Verlhac et al., 2014).

To gain more knowledge about the factors causing the variability in the results, we first tried to perform a cluster analysis using the Ward method. This grouped the participants into four clusters (see Fig. S8), with the main cluster (in yellow) including the 10 participants encompassing mainly the ones
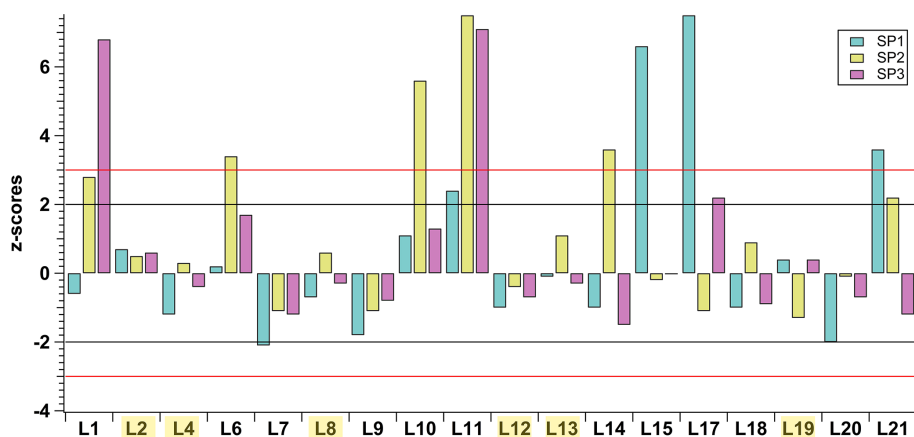
Figure 3. *Z* scores were calculated to evaluate each participant's performance in the interlaboratory comparison for each sample tested. Yellow-highlighted participants are the ones selected for the calculation of the assigned values. Horizontal black and red lines indicate boundaries for triggering an action signal as described in Sect. 2.7.1.

with satisfactory *z* scores. The clustering seemed independent of the instrument used and/or the time taken between the sample delivery and analysis (i.e. near the delivery time or later in February).

In a second step, a multiple linear regression model was run to evaluate the associations of the results obtained for the three samples, SP1, SP2 and SP3, considering a range of parameters, including the instrument used and the delivery and analysis time (Fig. 4, Table S2). The beta values are shown in Fig. 4, representing the association (effects) between the different parameters evaluated and the OP results obtained. In the model, the reference variables were the RI-URBANS DTT SOP and the results obtained with the plate-reader instrument. Regarding the instrument performance, the values provided by the cuvette-type spectrometer were higher than those obtained with plate readers in the case of SP1 and SP2 (showing significant overestimation in the case of SP2 p values < 0.05), while the results for SP3 were quite similar. In the case of LWCC, higher variability is observed when compared to both cuvette and plate reader for all the samples. SP1 LWCC results presented the highest variability, and SP2 results were significantly overestimated (at a 95 % confidence level) when compared with those obtained by the plate reader. The RI-URBANS SOP was adapted for plate readers and cuvettes in order to perform the measurements in similar conditions of concentrations for the reagents. This was not the case for the LWCC since we did not have all the necessary information concerning the specific devices used by participants. Figure 4 suggests that the specific conditions of the reaction are probably important factors for delivering an accurate value of OP. In Fig. S6, we showed that SP1's OP activity decreased over time during storage, but this ageing effect was not found to be significant in the model for either delivery or analysis time. The storage effect remained

consistent for SP2 and SP3, as there was a clear association between OP and delivery or analysis time (Fig. 4).

A ranking of the samples is also proposed to evaluate the OP activity of the samples tested in this ILC and its relative variability within the participants (only considering the results obtained with the simplified RI-URBANS protocol). For this purpose, SP1 was arbitrarily selected as the one with the highest OP activity with an assigned value of 100, and SP2 and SP3 were evaluated in function of SP1. Figure S12 shows the results obtained for the relative ranking of the samples. It can be noted that most of the participants presented similar relative variability with SP1 > SP2 > SP3. Some exceptions were observed for L1, L17 and L19, which obtained higher ranking for SP3 than SP2. A higher variability in the relative activity is obtained for SP2 than for SP3. Within the participants showing a higher relative ranking for SP2 (higher than 50 % compared to SP1), most of them used either cuvette-type or LWCC instruments (except L20), suggesting some overestimation in the results using these instruments. Overall, this similar ranking for the samples achieved by most of the groups is noticeable and very encouraging. In fact, most of the data treatment performed on OP with atmospheric variables or health data relies on associations and regressions where the relative variability of a time series is of utmost importance, more than the absolute value.

3.5 Comparison with other OP tests provided

Participants were also invited to report results obtained using other OP assays. Since not all participants submitted results from equivalent "home OP" tests, we exclusively focus on the outcomes obtained through the DTT-home protocols involving 13 participants (Table S3, Fig. S9). It is important to note that the RI-URBANS DTT protocol was simplified and does not include EDTA, Tris-HCl, TCA or Chelex®, whereas DTT-home protocols are diverse and should exhibit at least
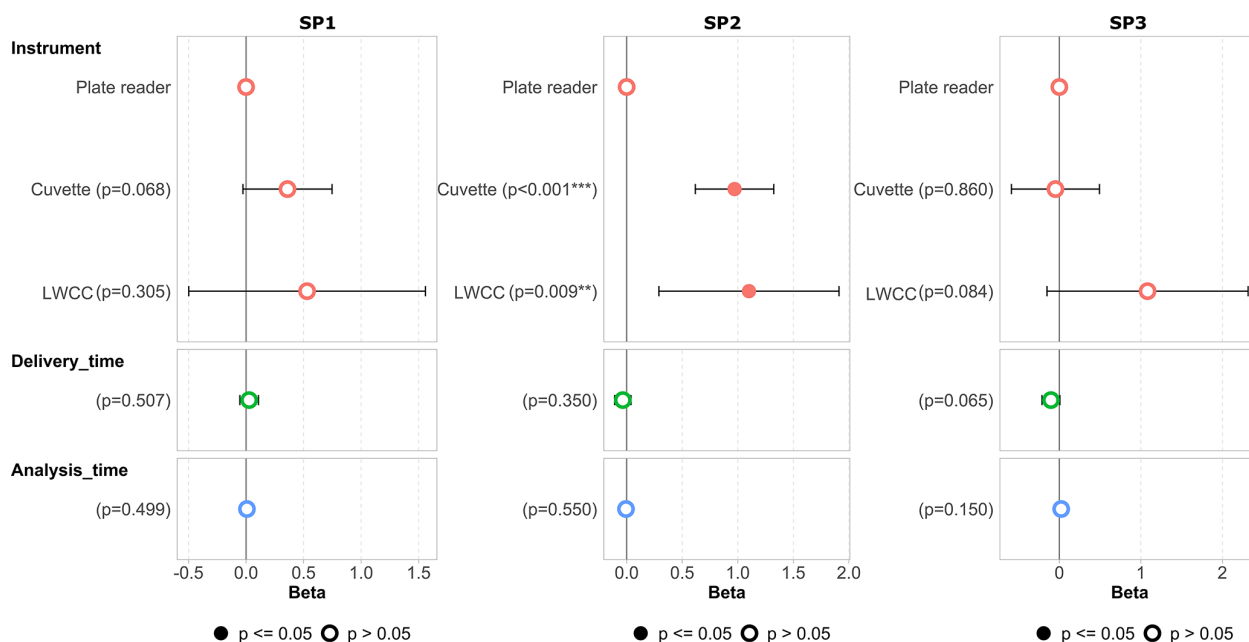
Figure 4. Associations (beta in $nmol\,min^{-1}\,\mu g^{-1}$) between OP DTT values for SP1, SP2 and SP3 using the RI-URBANS protocol and technical parameters, including the instrument used and the delivery and analysis time obtained by applying an adjusted multiple linear regression model. Coloured dots represent the results with p values ≤ 0.05, and white dots represent the results with p values > 0.05.

one of up to three of the last-mentioned compounds at different steps of the reaction; however, this is challenging to evaluate because not all groups have submitted the protocols related to their home results.

We have first evaluated the COV individually from the results obtained with protocols (RI-URBANS SOP versus all DTT-home) for each test sample. Figure 5 shows that lower COVs are generally observed in the performance of the DTT-home protocols (more details can be found in Table S3). However, 6 out of 12 participants presented similar COVs (within 20 %) for the two protocols. These results could indicate that the use of a simplified OP protocol needs some extent of training and guidance before its application. In addition, some of the participants presented higher COV values (L1 and L13 for DTT-home) when using the LWCC instrument. The lack of a simplified protocol for this instrument did not seem to be a major issue, as the application of the DTT-home protocols was also associated with high COV.

Another multiple linear regression model was run to evaluate the main differences in the results obtained between RI-URBANS SOP and DTT-home protocols (Fig. 6). For SP1, a significant overestimation was observed for the DTT-home protocol and the opposite was observed for SP3. In the case of SP2, there was no statistically significant difference between both protocols. More details on the concentration of DTT for each DTT method could provide more insight into this trend. Regarding the instrument performance, the LWCC presents poorer results compared to the cuvette and plate reader for all the samples, which is opposite to the

trend observed with the results of the RI-URBANS protocol only (Fig. 4). However, the results are significantly underestimated for SP3 only ($p < 0.05$). For the cuvette-based measurements, the results are higher than those obtained with plate readers in the case of SP1 and SP2 (significantly overestimated for SP2) and similar for SP3, which is in line with the direction observed with the RI-URBANS protocol only. The delivery and analysis time show a statistically significant lack of effect for SP2 (analysis time) and an underestimation for SP3 (delivery time) but nothing significant for SP1, although it had undergone ageing in the tests of the ILC organiser. Since the effects are very small compared to the effects of the protocol, or the instrument, these two variables (delivery and analysis time) may cause a greater impact on DTT values when protocols are harmonised but not to date.

The results obtained by the participants (the z-score evaluation) were also added to the former model to evaluate the effect on OP values while adjusting the protocol variables (Fig. S10). The results show a significant OP overestimation of all the samples for the labs with poor performances in SP2 and SP3 samples, as well as a significant underestimation in the OP value obtained for SP2 for the group with intermediate performance.

Hierarchical cluster analysis was conducted, incorporating both the RI-URBANS DTT and DTT-home outcomes (Fig. S11). Because some participants did not implement a DTT-home protocol, the cluster analysis involved a reduced set of OP values. The results reveal the presence of four primary clusters, with the predominant cluster encompassing
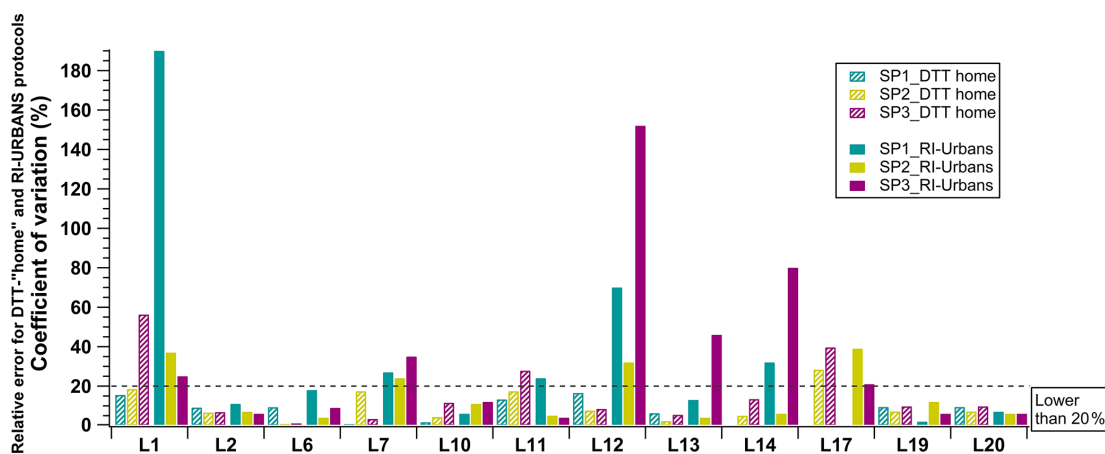
Figure 5. Coefficients of variation of each participant for the three samples tested from triplicates using both the RI-URBANS DTT and the DTT-home protocols. The average and standard deviation of each participant are also detailed in Table S3.
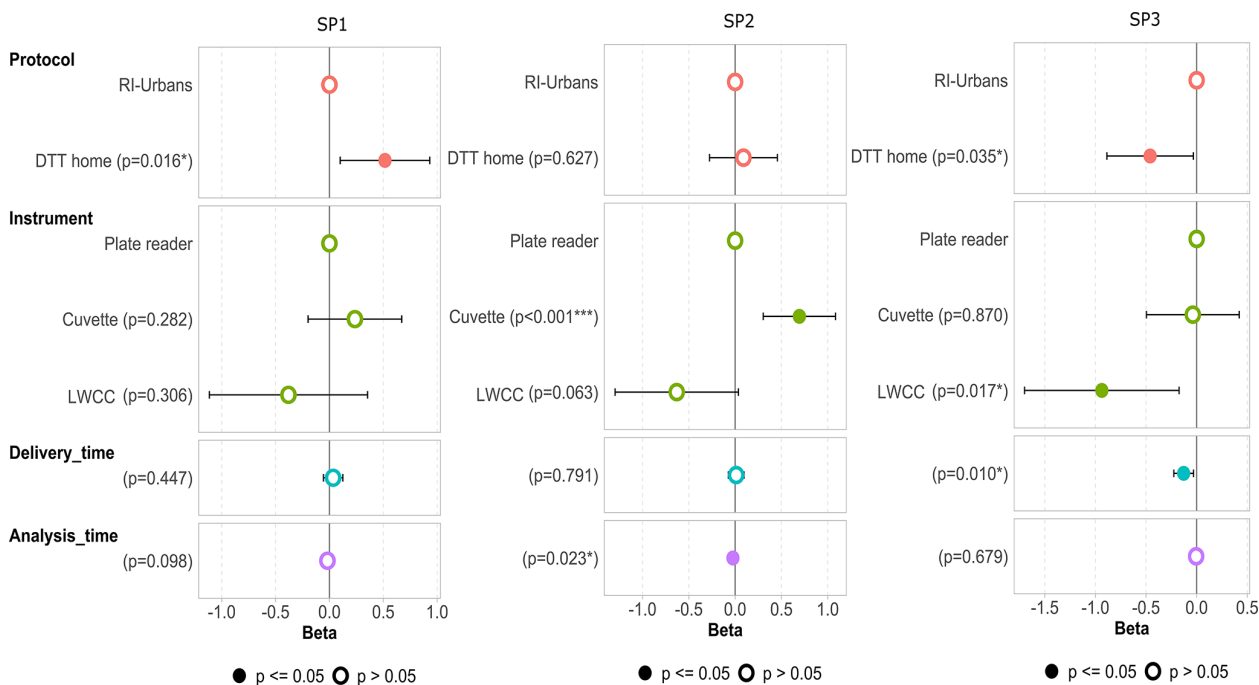

Figure 6. Associations (beta in $nmol\,min^{-1}\,\mu g^{-1}$) between OP DTT values obtained for SP1, SP2 and SP3 and the different parameters of the ILC, including the DTT protocol, the instrument used, and the delivery and analysis time obtained by applying a multiple linear model. Coloured dots represent the results with p values ≤ 0.05, and white dots represent the results with p values > 0.05.

most participants (8 out of 12 for DTT-home). The participants' results within the green main cluster largely align with those derived from the RI-URBANS DTT outcomes, encompassing the groups in the two primary clusters (Fig. S11). This assessment illustrates some consistency of results obtained across various OP DTT protocols. Some of the participants with more reliable results for the RI-URBANS DTT SOP maintain their consistency regardless of the protocol used. However, some of those that did not show an acceptable performance for the simplified protocol (i.e. L1 and L10) pre-

sented a better performance for DTT-home, and the opposite was observed for L19 (almost for SP1).

Finally, to assess the performance of the participants in the DTT-home protocols, a comparable approach to the simplified RI-URBANS SOP was employed for those participants who supplied OP results. The z scores were computed using the assigned values of each sample (SP1–SP3), obtained with the RI-URBANS SOP application. Figure 7 illustrates the z scores of the OP results obtained through the application of the DTT-home protocols, revealing a significant
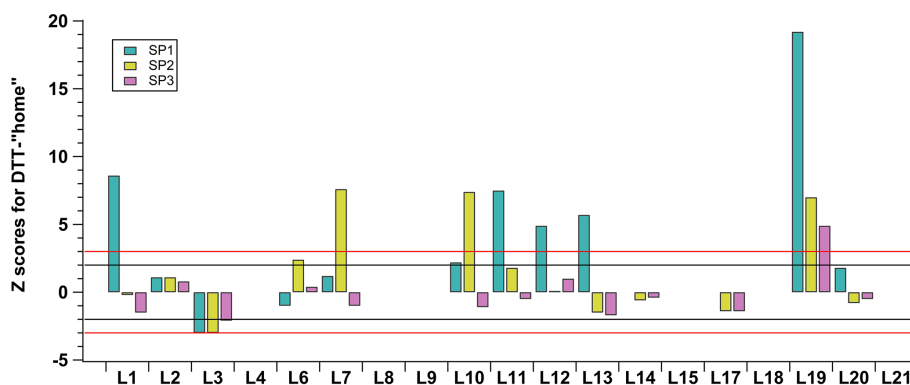
Figure 7. Z scores were calculated for the DTT-home protocol results to evaluate each participant's performance in reference to the RI-URBANS assigned (consensus) values for each sample tested. Horizontal black and red lines indicate boundaries for triggering an action signal as described in Sect. 2.7.1.

variation in the outcomes. Only five participants managed to produce satisfactory results for all the tested samples. Despite the fact that the COV of the participants using DTT-home protocols showed an improvement over the results of the simplified DTT SOP (Fig. 5), the outcomes are still distant from the consensus values of the samples obtained in this exercise. The results indicate a high degree of variability in the OP activity using home OP methodologies, underscoring the pressing requirement for standardised methods and harmonised protocols to ensure more reliable OP research.

4 Strengths and limitations of this first intercomparison

The greatest strength of this ILC was the high number of participants (20) enhancing the comprehensiveness and diversity of the study and allowing for a broader range of perspectives and expertise. This also allowed comprehensive collaborative discussions during the preparation phase, promoting knowledge exchange and consensus-building, as well as contributing to a more robust ILC design. These all show a willingness from the groups to be actively part of the development of the intercomparison and to pursue a harmonisation of the OP measurements.

The development of the first simplified OP DTT protocol (available in Sect. S1) also consolidates the experimental experience of the participants, fostering methodological consistency across different research groups as a first step toward method harmonisation. Finally, the collaboration with the JRC, an independent organisation for the assessment of results, adds credibility and objectivity to the study, ensuring that findings are impartially evaluated.

The sharing of liquid samples in this comparison comes with both advantages and limitations. On one hand, it eliminates biases associated with extraction methods and solvent purity. However, some samples exhibited signs of ageing during the interlaboratory comparison duration (though this was

not identified as a critical parameter when identifying the main causes of variability). In addition, this approach introduced certain challenges with some of the home OP protocols which were designed originally to be used with solid samples. Finally, all liquid extracts should be provided with a similar PM concentration to limit known non-linearities and to avoid potential saturation issues, as can be the case with LWCC instruments. Next, ILCs should, in the future, include the whole chain assessment, including the extraction step.

The testing of three samples with different patterns makes it difficult to draw unilateral conclusions. A larger sample size could support a more robust statistical analysis of the results, particularly for the factors determining OP variability. Moreover, the inclusion of samples that are readily accessible worldwide, such as standard reference materials, could facilitate the future adaptation of a simplified OP DTT protocol while allowing comparison with an assigned (consensus) value.

The sole focus on the DTT method for the ILC could limit the broader evaluation of OP. The addition of other OP assays, such as those included in previous inter-comparison studies (Ayres et al., 2008; Calas et al., 2018; Shahpoury et al., 2022), could provide a more comprehensive understanding of the performances of the different groups involved in the OP domain of research. There are contrasting reports about the relative sensitivity of the DTT assay to various organic and inorganic PM components, with some studies showing higher reactivity towards the organic fraction. Therefore, additional consensus studies would be needed to assess this aspect and the comparability of DTT to other OP metrics that rely on proper lung antioxidants and could be considered more physiologically relevant. Such studies could support the identification of chemical species which should be prioritised for future air quality management programmes.

In conclusion, while this ILC of OP has highlighted considerable variability in the performance of the assay between

groups, it has notable strengths and provides a starting point towards the harmonisation of OP measurements.

5   Recommendations for standardisation of OP protocols

Based on the findings of this ILC and also on the general literature about OP, some recommendations for further standardising OP DTT measurements are proposed (Table 1). These include guidelines for sample and laboratory conditions, instrument type and calibration, and the reporting of results. Additionally, a reference material (1,4-naphthoquinone, copper or other solution at a known concentration) should be proposed to facilitate future ILCs. Additionally, since some differences were observed in the results obtained from the use of the simplified SOP compared with the "home-developed" DTT protocols, the harmonisation of procedures is needed to ensure data comparability. We describe below the crucial parameters that need to be considered in the move toward greater harmonisation of OP methodologies.

OP assay selection

– To date, it remains unclear which oxidative potential (OP) assay is most effective at predicting health outcomes related to oxidative stress. This uncertainty arises because different assays yield markedly different results for the same particulate matter (PM) samples. Additionally, even OP values from the same assay are often linked to various PM components and sources, depending on the different studies (He and Zhang, 2022).

   Thus, based on current knowledge and epidemiological evidence, two complementary OP assays (a thiol-based probe (OP DTT or OPGSH) and another one (OPOH, OPAA or another)) are recommended to provide a better picture of the potential oxidising damage from PM compounds and to strengthen the power of epidemiological studies. These aspects were previously discussed in a recent work integrating five different OP assays (Dominutti et al., 2023). Finally, the final choice of the best OP test (or combination) must be based on epidemiological evidence, which has begun to be studied only recently and needs more hindsight to be determined.

Sampling

– OP can be analysed in filter samples conventionally collected for air quality monitoring using small portions of these if adequately preserved (frozen). Pre-burn quartz filters or Teflon filters are appropriate, and blank filters must be measured to remove the background induced by the matrix of the material. A previous study had shown no differences in the OP DTT values observed using Quartz or Teflon filters (Frezzini et al., 2022). However, it should be further evaluated when other OP assays are considered.

Sample storage

– Previous studies that evaluated the effect of storage time and conditions did not show a substantial effect on the OP DTT results (Frezzini et al., 2022). However, we recommend that PM samples should be immediately transported to the lab after sampling. The filters must be kept cold after sampling (at 4 °C if the OP analysis is done within a few days after collection or −20 °C if the analysis is delayed).

– The lifetime of the ROS may be very short, and measurements of OP on PM-extracted filters are likely affected both by the age of the samples, how they have been sampled and stored, and the nature of the extraction methodology. How all of these processes have an impact on the ageing of samples and the ultimate quantification of OP needs to be addressed. Ageing studies should be performed for each OP assay in the long term to define the maximum storage time of aerosol filters at low versus ambient temperature conditions.

   Some OP components might be so short-lived that only online techniques are warranted.

Laboratory conditions

– OP assays are "trace" detection assays that require clean ambient conditions and high-quality reagents free of metal contamination. Considerations should be given to the use of certified clean rooms or proper laminar flow bench stations to prevent contamination of the samples.

– Regarding the use of clean material, vials, cones and spatulas have to be washed before use (5 % HNO$_3$ bath to remove metals and rinsed three times in ultra-pure water before drying in laminar flow).

– Control laboratory temperatures and light exposure by using dark polypropylene tubes at least for reactants

Extraction step

– The extraction step may be highly variable according to the procedures used, and several parameters are known to have an impact on OP results, such as the choice of the solvent, the concentration of buffer, the way of agitation and the quantification of the final extracted mass. Notably, the ultrasonication of PM samples in aqueous solutions generates ROS (Miljevic et al., 2014), and it could introduce artefacts in OP measurement. This effect was also observed in the work of Frezzini et al. (2022), where different extraction methods were evaluated, with ultrasonic baths overestimating the results observed.

   The effect of the solvent used was not evaluated in this ILC exercise. However, we recommend the use of ultra-pure water or simulated lining fluid for the sample ex-

Table 1. Summary of recommendations for future OP measurements on filters.

Condition or step	Recommendations
OP assay selection	– Two complementary OP assays (a thiol-based probe (OP DTT or OPGSH) and another one (OPOH (ROS quantification through OH radical), OPAA or others)) are recommended to provide a better picture of the potential oxidising damage from PM compounds.
Sampling	– Pre-burn quartz filters or Teflon filters are appropriate, and blank filters must be measured to remove the background induced by the matrix of the material.
Samples storage	– The PM filters must be kept cold after sampling (at 4 °C if the OP analysis is done within a few days after collection or −18 °C or −20 °C if the analysis is delayed).
Laboratory conditions	– Clean conditions (including certified clean rooms or proper laminar flow bench stations). – High-quality reagents free of metal contamination. – Use of clean material, which must be washed before use (5 % HNO$_3$ bath to remove metals, rinsed three times in ultra-pure water and dried in laminar flow bench stations). – Control laboratory temperatures and light exposure by using dark polypropylene tubes.
Reaction step	– Several aspects in the reaction process affect the OP value, and to minimise their impact, standard conditions should be fixed as the initial concentration of reactants, ratio of reactant/sample, time of reaction, the temperature of the reaction (37 °C), agitation (mixing samples) and the type of measurements (kinetic or end-point value).
Instrument calibration	– Investigate the optimal frequency for the calibration of spectrophotometers for such assays.

traction. Future ILC exercises should include the evaluation on the extraction conditions, including solvent use and methods.

Reaction step

– Several aspects in the reaction process affect the OP value, like the initial concentration of reactants (since the DTT test is mass-dependent; Charrier et al., 2016), ratio of reactant/sample, time of reaction (some compounds present a non-linear reaction over time), the temperature of the reaction (which should be standardised to 37 °C), agitation (mixing samples) and the type of measurements (kinetic or end-point value), etc.

– The current literature mainly addresses the extraction or reaction parameters separately. We advise that the whole chain factors should be evaluated together to quantify their relative impact on the results.

Development of a reference material with a certified "OP value"

– The set-up of reference material or in-house standard solutions (in collaboration with reference institutions JRC or NIST, for instance) with a known OP value could help laboratories test and train themselves on the OP protocol before testing the unknown ILC samples. This is something to be developed and tested in future ILCs.

Instrument calibration

– Investigate the optimal frequency for the calibration of spectrophotometers for such assays.

Report of results and units

– The calculation of OP DTT activity during this ILC involved a conversion using a calibration curve. Since the OP activity measures the rate of a chemical reaction and not a concentration, for future comparison exercises, the possibility of exploring alternative methods for OP calculation should be tested. To date, results are mass-normalised in nmol antioxidant min^{-1} µg^{-1} or volume-normalised in nmol antioxidant min^{-1} m^{-3}. The OP per microgram refers to the reactivity of 1 µg of the tested PM, whereas the OP per cubic metre refers to the exposure of 1 m^3 of inhaled air.

6 Conclusions

This study represents an innovative effort as the first interlaboratory OP exercise specifically aimed at harmonising this OP assay. This exercise provides the very first roadmap for refining interlaboratory comparisons of OP, fostering greater confidence in the reliability of OP data and encouraging the scientific community to advance towards global OP harmonisation.

This first exercise focused on OP DTT, as it is widely used within the scientific community and has already shown pos-

itive associations with health outcomes (Bates et al., 2015; Borlaza et al., 2022; Dabass et al., 2018; Donaldson et al., 2001; Gao et al., 2020; Marsal et al., 2023; Weichenthal et al., 2016b, a, c). Even if there are several crucial points to be evaluated and harmonised in the whole chain of the determination of OP (sampling methods, sample storage, extraction conditions and methods) as well as the use of different OP assays, this first ILC engaging several research laboratories paves the way for future developments towards the standardisation of OP methods.

Our findings emphasise both the strengths and challenges associated with the use of the current OP DTT assay for driving a measurement of PM OP. Overall, half of the participants achieved results falling within a satisfactory range of z scores for this test. The participating group performance levels are comparable to those observed in initial ILCs for PAHs in the 2010s (Grandesso et al., 2012; Verlhac et al., 2014). While notable agreement was observed in certain samples and between several groups, discrepancies and variability were also identified, emphasising the need for harmonisation in the procedures and conditions. A number of factors may contribute to the underperformance observed in certain samples and participants. The main reasons are not clear, but the analysis conditions in the participating laboratories and the lack of experience in this type of metrological exercise are possible causes. The standardisation of protocols and harmonisation of procedures emerged as critical components to ensure the accuracy and comparability of OP data across laboratories. This collaborative approach fosters a more robust OP science, facilitates data exchange and integration, and will ultimately contribute to a better understanding of the health impacts associated with PM exposure, allowing for more accurate exposure assessments and regulatory decisions.

Author contributions. GU designed the study and supervised the OP tests for the development of the simplified RI-URBANS DTT protocol. CR, TM, RE and PAD tested the different steps and prepared the samples and logistics of the intercomparison exercise. GH, RMH, AN, IM, AB, NM and GU evaluated and decided on the samples to be compared and the protocol content. JPP and FC performed the data analysis and evaluated the performance of each group. PAD and AM developed the multiple regression model. PAD processed the data of the study and wrote the paper together with JLJ and GU before a first review by the RI-URBANS members. All authors participated in the interlaboratory comparison. All authors reviewed and edited the manuscript.

Review statement. This paper was edited by Pierre Herckes and reviewed by four anonymous referees.

References

Alkoussa, S., Hulo, S., Courcot, D., Billet, S., and Martin, P. J.: Extracellular vesicles as actors in the air pollution related cardiopulmonary diseases, Crit. Rev. Toxicol., 50, 402–423, https://doi.org/10.1080/10408444.2020.1763252, 2020.

Ayres, J. G., Borm, P., Cassee, F. R., Castranova, V., Donaldson, K., Ghio, A., Harrison, R. M., Hider, R., Kelly, F., Kooter, I. M., Marano, F., Maynard, R. L., Mudway, I., Nel, A., Sioutas, C., Smith, S., Baeza-Squiban, A., Cho, A., Duggan, S., and Froines, J.: Evaluating the toxicity of airborne particulate matter and nanoparticles by measuring oxidative stress potential – A workshop report and consensus statement, Inhal. Toxicol., 20, 75–99, https://doi.org/10.1080/08958370701665517, 2008.

Bates, J. T., Weber, R. J., Abrams, J., Verma, V., Fang, T., Klein, M., Strickland, M. J., Sarnat, S. E., Chang, H. H., Mulholland, J. A., Tolbert, P. E., and Russell, A. G.: Reactive Oxygen Species Generation Linked to Sources of Atmospheric Particulate Matter and Cardiorespiratory Effects, Environ. Sci. Technol., 49, 13605–13612, https://doi.org/10.1021/acs.est.5b02967, 2015.

Bates, J. T., Fang, T., Verma, V., Zeng, L., Weber, R. J., Tolbert, P. E., Abrams, J. Y., Sarnat, S. E., Klein, M., Mulholland, J. A., and Russell, A. G.: Review of Acellular Assays of Ambient Particulate Matter Oxidative Potential: Methods and Relationships with Composition, Sources, and Health Effects, Environ. Sci. Technol., 53, 4003–4019, https://doi.org/10.1021/acs.est.8b03430, 2019.

Bhattu, D., Tripathi, S. N., Bhowmik, H. S., Moschos, V., Lee, C. P., Rauber, M., Salazar, G., Abbaszade, G., Cui, T., Slowik, J. G., Vats, P., Mishra, S., Lalchandani, V., Satish, R., Rai, P., Casotto, R., Tobler, A., Kumar, V., Hao, Y., Qi, L., Khare, P., Manousakas, M. I., Wang, Q., Han, Y., Tian, J., Darfeuil, S., Minguillon, M. C., Hueglin, C., Conil, S., Rastogi, N., Srivastava, A. K., Ganguly, D., Bjelic, S., Canonaco, F., Schnelle-Kreis, J., Dominutti, P. A., Jaffrezo, J.-L., Szidat, S., Chen, Y., Cao, J., Baltensperger, U., Uzu, G., Daellenbach, K. R., El Haddad, I., and Prévôt, A. S. H.: Local incomplete combustion emissions define the PM$_{2.5}$ oxidative potential in Northern India, Nat. Commun., 15, 3517, https://doi.org/10.1038/s41467-024-47785-5, 2024.

Borlaza, L. J. S., Uzu, G., Ouidir, M., Lyon-Caen, S., Marsal, A., Weber, S., Siroux, V., Lepeule, J., Boudier, A., Jaffrezo, J.-L., Slama, R., Lyon-Caen, S., Siroux, V., Lepeule, J., Philippat, C., Slama, R., Hofmann, P., Hullo, E., Llerena, C., Quentin, J., Pin, I., Eyriey, E., Licinia, A., Vellement, A., Morin, X., Morlot, A., and the SEPAGES cohort study group: Personal exposure to PM$_{2.5}$ oxidative potential and its association to birth outcomes, J. Expo. Sci. Environ. Epidemiol., 33, 416–426, https://doi.org/10.1038/s41370-022-00487-w, 2022.

Calas, A., Uzu, G., Martins, J. M. F., Voisin, D., Spadini, L., Lacroix, T., and Jaffrezo, J.-L.: The importance of simulated lung fluid (SLF) extractions for a more relevant evaluation of the oxidative potential of particulate matter, Sci. Rep., 7, 11617, https://doi.org/10.1038/s41598-017-11979-3, 2017.

Calas, A., Uzu, G., Kelly, F. J., Houdier, S., Martins, J. M. F., Thomas, F., Molton, F., Charron, A., Dunster, C., Oliete, A., Jacob, V., Besombes, J.-L., Chevrier, F., and Jaffrezo, J.-L.: Comparison between five acellular oxidative potential measurement assays performed with detailed chemistry on PM$_{10}$ samples from the city of Chamonix (France), Atmos. Chem. Phys., 18, 7863–7875, https://doi.org/10.5194/acp-18-7863-2018, 2018.

Cassee, F. R., Héroux, M.-E., Gerlofs-Nijland, M. E., and Kelly, F. J.: Particulate matter beyond mass: recent health evidence on the role of fractions, chemical constituents and sources of emission, Inhal. Toxicol., 25, 802–812, https://doi.org/10.3109/08958378.2013.850127, 2013.

Charrier, J. G. and Anastasio, C.: On dithiothreitol (DTT) as a measure of oxidative potential for ambient particles: evidence for the importance of soluble transition metals, Atmos. Chem. Phys., 12, 9321–9333, https://doi.org/10.5194/acp-12-9321-2012, 2012.

Charrier, J. G., Richards-Henderson, N. K., Bein, K. J., McFall, A. S., Wexler, A. S., and Anastasio, C.: Oxidant production from source-oriented particulate matter – Part 1: Oxidative potential using the dithiothreitol (DTT) assay, Atmos. Chem. Phys., 15, 2327–2340, https://doi.org/10.5194/acp-15-2327-2015, 2015.

Charrier, J. G., McFall, A. S., Vu, K. K.-T., Baroi, J., Olea, C., Hasson, A., and Anastasio, C.: A bias in the "mass-normalized" DTT response – An effect of non-linear concentration-response curves for copper and manganese, Atmos. Environ., 144, 325–334, https://doi.org/10.1016/j.atmosenv.2016.08.071, 2016.

Cho, A. K., Sioutas, C., Miguel, A. H., Kumagai, Y., Schmitz, D. A., Singh, M., Eiguren-Fernandez, A., and Froines, J. R.: Redox activity of airborne particulate matter at different sites in the Los Angeles Basin, Environ. Res., 99, 40–47, https://doi.org/10.1016/j.envres.2005.01.003, 2005.

Council of the European Union: Directive (EU) 2024/2881 of the European Parliament and of the Council of 23 October 2024 on ambient air quality and cleaner air for Europe (recast), Document 32024L2881, PE/88/2024/REV/1 OJ L, 2024/2881, http://data.europa.eu/eli/dir/2024/2881/oj (last access: 20 December 2024), 2024.

Dabass, A., Talbott, E. O., Rager, J. R., Marsh, G. M., Venkat, A., Holguin, F., and Sharma, R. K.: Systemic inflammatory markers associated with cardiovascular disease and acute and chronic exposure to fine particulate matter air pollution (PM$_{2.5}$) among US NHANES adults with metabolic syndrome, Environ. Res., 161, 485–491, https://doi.org/10.1016/j.envres.2017.11.042, 2018.

Daellenbach, K. R., Stefenelli, G., Bozzetti, C., Vlachou, A., Fermo, P., Gonzalez, R., Piazzalunga, A., Colombi, C., Canonaco, F., Hueglin, C., Kasper-Giebl, A., Jaffrezo, J.-L., Bianchi, F., Slowik, J. G., Baltensperger, U., El-Haddad, I., and Prévôt, A. S. H.: Long-term chemical analysis and organic aerosol source apportionment at nine sites in central Europe: source identification and uncertainty assessment, Atmos. Chem. Phys., 17, 13265–13282, https://doi.org/10.5194/acp-17-13265-2017, 2017.

Danehy, J. P., Elia, V. J., and Lavelle, C. J.: Alkaline decomposition of organic disulfides. IV. Limitation on the use of Ellman's reagent. 2,2'-Dinitro-5,5'-dithiodibenzoic acid, J. Org. Chem., 36, 1003–1005, https://doi.org/10.1021/jo00806a036, 1971.

Dominutti, P. A., Borlaza, L. J. S., Sauvain, J.-J., Ngoc Thuy, V. D., Houdier, S., Suarez, G., Jaffrezo, J.-L., Tobin, S., Trébuchon, C., Socquet, S., Moussu, E., Mary, G., and Uzu, G.: Source apportionment of oxidative potential depends on the choice of the assay: insights into 5 protocols comparison and implications

for mitigation measures, Environ. Sci. Atmos., 3, 1497–1512, https://doi.org/10.1039/D3EA00007A, 2023.

Donaldson, K., Stone, V., Seaton, A., and MacNee, W.: Ambient particle inhalation and the cardiovascular system: Potential mechanisms, Environ. Health Perspect., 109, 523–527, https://doi.org/10.1289/ehp.01109s4523, 2001.

Frezzini, M. A., De Francesco, N., Massimi, L., and Canepari, S.: Effects of operating conditions on PM oxidative potential assays, Atmos. Environ., 268, 118802, https://doi.org/10.1016/j.atmosenv.2021.118802, 2022.

Gao, D., Mulholland, J. A., Russell, A. G., and Weber, R. J.: Characterization of water-insoluble oxidative potential of $PM_{2.5}$ using the dithiothreitol assay, Atmos. Environ., 224, 117327, https://doi.org/10.1016/j.atmosenv.2020.117327, 2020.

Gao, Y., Wang, P., Chu, Y., Kang, F., Cheng, Y., Repo, E., Feng, M., Yu, X., and Zeng, H.: Redox property of coordinated iron ion enables activation of O_2 via in-situ generated H_2O_2 and additionally added H_2O_2 in EDTA-chelated Fenton reaction, Water Res., 248, 120826, https://doi.org/10.1016/j.watres.2023.120826, 2024.

Grandesso, E., Kowalewski, K., and Perez Ballesta, P.: First EC-JRC PAHs inter-laboratory comparison on PM_{10} quartz filters, European Commission, Joint Research Centre, Institute for Environment and Sustainability, Luxembourg, ISBN 978-92-79-22782-0, https://doi.org/10.2788/63064, 2012.

Guascito, M. R., Lionetto, M. G., Mazzotta, F., Conte, M., Giordano, M. E., Caricato, R., De Bartolomeo, A. R., Dinoi, A., Cesari, D., Merico, E., Mazzotta, L., and Contini, D.: Characterisation of the correlations between oxidative potential and in vitro biological effects of PM_{10} at three sites in the central Mediterranean, J. Hazard. Mater., 448, 130872, https://doi.org/10.1016/j.jhazmat.2023.130872, 2023.

Hart, J. E., Liao, X., Hong, B., Puett, R. C., Yanosky, J. D., Suh, H., Kioumourtzoglou, M.-A., Spiegelman, D., and Laden, F.: The association of long-term exposure to $PM_{2.5}$ on all-cause mortality in the Nurses' Health Study and the impact of measurement-error correction, Environ. Health, 14, 38, https://doi.org/10.1186/s12940-015-0027-6, 2015.

He, L. and Zhang, J.: Particulate matter (PM) oxidative potential: Measurement methods and links to PM physicochemical characteristics and health effects, Crit. Rev. Environ. Sci. Technol., 53, 177–197, https://doi.org/10.1080/10643389.2022.2050148, 2022.

Huang, H., Yang, L., Liu, Y., Dong, G.-H., Chen, L., Li, S., Guo, Y., Xie, B., and Chen, G.: Long-term ambient air pollution exposure and DNA methylation of peripheral brain-derived neurotrophic factor promoter, Ecotoxicol. Environ. Saf., 244, 114061, https://doi.org/10.1016/j.ecoenv.2022.114061, 2022.

Kelly, F. J. and Fussell, J. C.: Role of oxidative stress in cardiovascular disease outcomes following exposure to ambient air pollution, Free Radic. Biol. Med., 110, 345–367, https://doi.org/10.1016/j.freeradbiomed.2017.06.019, 2017.

Kumagai, Y., Koide, S., Taguchi, K., Endo, A., Nakai, Y., Yoshikawa, T., and Shimojo, N.: Oxidation of proximal protein sulfhydryls by phenanthraquinone, a component of diesel exhaust particles, Chem. Res. Toxicol., 15, 483–489, https://doi.org/10.1021/tx0100993, 2002.

Laden, F., Schwartz, J., Speizer, F. E., and Dockery, D. W.: Reduction in Fine Particulate Air Pollution and Mor-tality, Am. J. Respir. Crit. Care Med., 173, 667–672, https://doi.org/10.1164/rccm.200503-443OC, 2006.

Leni, Z., Cassagnes, L. E., Daellenbach, K. R., Haddad, I. E., Vlachou, A., Uzu, G., Prévôt, A. S. H., Jaffrezo, J. L., Baumlin, N., Salathe, M., Baltensperger, U., Dommen, J., and Geiser, M.: Oxidative stress-induced inflammation in susceptible airways by anthropogenic aerosol, PLoS ONE, 15, 1–17, https://doi.org/10.1371/journal.pone.0233425, 2020.

Lepeule, J., Laden, F., Dockery, D., and Schwartz, J.: Chronic Exposure to Fine Particles and Mortality: An Extended Follow-up of the Harvard Six Cities Study from 1974 to 2009, Environ. Health Perspect., 120, 965–970, https://doi.org/10.1289/ehp.1104660, 2012.

Li, N., Sioutas, C., Cho, A., Schmitz, D., Misra, C., Sempf, J., Wang, M., Oberley, T., Froines, J., and Nel, A.: Ultra-fine particulate pollutants induce oxidative stress and mitochondrial damage, Environ. Health Perspect., 111, 455–460, https://doi.org/10.1289/ehp.6000, 2003.

Li, N., Xia, T., and Nel, A. E.: The role of oxidative stress in ambient particulate matter-induced lung diseases and its implications in the toxicity of engineered nanoparticles, Free Radic. Biol. Med., 44, 1689–1699, https://doi.org/10.1016/j.freeradbiomed.2008.01.028, 2008.

Li, Q., Wyatt, A., and Kamens, R. M.: Oxidant generation and toxicity enhancement of aged-diesel exhaust, Atmos. Environ., 43, 1037–1042, https://doi.org/10.1016/j.atmosenv.2008.11.018, 2009.

Marsal, A., Slama, R., Lyon-Caen, S., Borlaza, L. J. S., Jaffrezo, J.-L., Boudier, A., Darfeuil, S., Elazzouzi, R., Gioria, Y., Lepeule, J., Chartier, R., Pin, I., Quentin, J., Bayat, S., Uzu, G., Siroux, V., and the SEPAGES cohort study group: Prenatal Exposure to $PM_{2.5}$ Oxidative Potential and Lung Function in Infants and Preschool-Age Children: A Prospective Study, Environ. Health Perspect., 131, 017004, https://doi.org/10.1289/EHP11155, 2023.

Miljevic, B., Hedayat, F., Stevanovic, S., Fairfull-Smith, K. E., Bottle, S. E., and Ristovski, Z. D.: To Sonicate or Not to Sonicate PM Filters: Reactive Oxygen Species Generation Upon Ultrasonic Irradiation, Aerosol Sci. Tech., 48, 1276–1284, https://doi.org/10.1080/02786826.2014.981330, 2014.

Molina, C., Andrade, C., Manzano, C. A., Richard Toro, A., Verma, V., and Leiva-Guzmán, M. A.: Dithiothreitol-based oxidative potential for airborne particulate matter: an estimation of the associated uncertainty, Environ. Sci. Pollut. Res., 27, 29672–29680, https://doi.org/10.1007/s11356-020-09508-3, 2020.

Nicholson, S., Baccarelli, A., and Prada, D.: Role of brain extracellular vesicles in air pollution-related cognitive impairment and neurodegeneration, Environ. Res., 204, 112316, https://doi.org/10.1016/j.envres.2021.112316, 2022.

Park, M., Joo, H. S., Lee, K., Jang, M., Kim, S. D., Kim, I., Borlaza, L. J. S., Lim, H., Shin, H., Chung, K. H., Choi, Y.-H., Park, S. G., Bae, M.-S., Lee, J., Song, H., and Park, K.: Differential toxicities of fine particulate matters from various sources, Sci. Rep., 8, 17007, https://doi.org/10.1038/s41598-018-35398-0, 2018.

Sauvain, J. J., Deslarzes, S., and Riediker, M.: Nanoparticle reactivity toward dithiothreitol, Nanotoxicology, 2, 121–129, https://doi.org/10.1080/17435390802245716, 2008.

Shahpoury, P., Harner, T., Lammel, G., Lelieveld, S., Tong, H., and Wilson, J.: Development of an antioxidant assay to

study oxidative potential of airborne particulate matter, Atmos. Meas. Tech., 12, 6529–6539, https://doi.org/10.5194/amt-12-6529-2019, 2019.

Shahpoury, P., Zhang, Z. W., Filippi, A., Hildmann, S., Lelieveld, S., Mashtakov, B., Patel, B. R., Traub, A., Umbrio, D., Wietzoreck, M., Wilson, J., Berkemeier, T., Celo, V., Dabek-Zlotorzynska, E., Evans, G., Harner, T., Kerman, K., Lammel, G., Noroozifar, M., Pöschl, U., and Tong, H.: Inter-comparison of oxidative potential metrics for airborne particles identifies differences between acellular chemical assays, Atmos. Pollut. Res., 13, 101596, https://doi.org/10.1016/j.apr.2022.101596, 2022.

Shankar, K. and Mehendale, H. M.: Oxidative Stress, 3rd edn., Elsevier, 735–737, https://doi.org/10.1016/B978-0-12-386454-3.00345-6, 2014.

Sies, H.: On the history of oxidative stress: Concept and some aspects of current development, Curr. Opin. Toxicol., 7, 122–126, https://doi.org/10.1016/j.cotox.2018.01.002, 2018.

Thbayh, D. K., Palusiak, M., Viskolcz, B., and Fiser, B.: Comparative study of the antioxidant capability of EDTA and Irganox, Heliyon, 9, e16064, https://doi.org/10.1016/j.heliyon.2023.e16064, 2023.

Uzu, G., Sauvain, J. J., Baeza-Squiban, A., Riediker, M., Hohl, M. S. S., Val, S., Tack, K., Denys, S., Pradère, P., and Dumat, C.: In vitro assessment of the pulmonary toxicity and gastric availability of lead-rich particles from a lead recycling plant, Environ. Sci. Technol., 45, 7888–7895, https://doi.org/10.1021/es200374c, 2011.

Verlhac, S., Leoz, E., Albinet, A., Cabillic, J., Lalere, B., and Fallot, C.: Comparaison interlaboratoires sur l'analyse des HAP dans l'air ambiant, Laboratoire Central de Surveillance de la Qualité de l'Air, DRC-14-144350-09750A, 2014.

Weichenthal, S., Lavigne, E., Evans, G., Pollitt, K., and Burnett, R. T.: Ambient $PM_{2.5}$ and risk of emergency room visits for myocardial infarction: impact of regional $PM_{2.5}$ oxidative potential: a case-crossover study, Environ. Health, 15, 46, https://doi.org/10.1186/s12940-016-0129-9, 2016a.

Weichenthal, S., Crouse, D. L., Pinault, L., Godri-Pollitt, K., Lavigne, E., Evans, G., van Donkelaar, A., Martin, R. V., and Burnett, R. T.: Oxidative burden of fine particulate air pollution and risk of cause-specific mortality in the Canadian Census Health and Environment Cohort (CanCHEC), Environ. Res., 146, 92–99, https://doi.org/10.1016/j.envres.2015.12.013, 2016b.

Weichenthal, S. A., Lavigne, E., Evans, G. J., Godri Pollitt, K. J., and Burnett, R. T.: Fine particulate matter and emergency room visits for respiratory illness: Effect modification by oxidative potential, Am. J. Respir. Crit. Care Med., 194, 577–586, https://doi.org/10.1164/rccm.201512-2434OC, 2016c.

Weichenthal, S., Shekarrizfard, M., Traub, A., Kulka, R., Al-Rijleh, K., Anowar, S., Evans, G., and Hatzopoulou, M.: Within-City Spatial Variations in Multiple Measures of $PM_{2.5}$ Oxidative Potential in Toronto, Canada, Environ. Sci. Technol., 53, 2799–2810, https://doi.org/10.1021/acs.est.8b05543, 2019.

WHO: Review of evidence on health aspects of air pollution REVIHAAP Project – Technical Report, World Health Organization, Regional Office for Europe, WHO/EURO:2013-4101-43860-61757, 2013.

WHO: WHO Global air quality guidelines. Particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide, World Health Organization, ISBN 978-92-4-003422-8, 2021.

WHO: Sustainable Development Goal indicator 3.9.1: mortality attributed to air pollution, World Health Organization, ISBN 978-92-4009-914-2, 2024.

Wilker, E. H., Osman, M., and Weisskopf, M. G.: Ambient air pollution and clinical dementia: systematic review and meta-analysis, BMJ, 381, e071620, https://doi.org/10.1136/bmj-2022-071620, 2023.

Zare Sakhvidi, M. J., Yang, J., Lequy, E., Chen, J., De Hoogh, K., Letellier, N., Mortamais, M., Ozguler, A., Vienneau, D., Zins, M., Goldberg, M., Berr, C., and Jacquemin, B.: Outdoor air pollution exposure and cognitive performance: findings from the enrolment phase of the CONSTANCES cohort, Lancet Planet. Health, 6, e219–e229, https://doi.org/10.1016/S2542-5196(22)00001-8, 2022.