Atmospheric Environment 352 (2025) 121219

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Comparison of NO_2 diffusion tube measurement methods and related uncertainties

David Medland [a,*], Gabriel Garcia [a], Tom Gardiner [a], Nicholas A. Martin [a], Valerio Ferracci [a], Ashley Wilkins [a], David Fryer [a], Tom Holmes [a], Pete Edwards [b], Sebastian Diez [b], David Butterfield [a]

[a] *Atmospheric Environmental Science Department, National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, United Kingdom*
[b] *University of York, United Kingdom*

HIGHLIGHTS

- Unsheltered, modified tubes perform better than conventional Palmes diffusion tubes.
- Palmes diffusion tubes deliver better performance when deployed in shelters.
- Palmes diffusion tube uncertainty is dominated by wind speed effects.
- Uncertainty assessment identifies key uncertainty sources over different timescales.

ARTICLE INFO

Keywords:
Diffusive sampling
Nitrogen dioxide (NO_2)
Measurement uncertainty
Local air quality management
Diffusion tubes

ABSTRACT

Conventional Palmes Diffusion Tubes (PDTs) are extensively employed by UK Local Authorities for measuring NO_2 in air quality monitoring studies. These devices are known to suffer from biases due to from the effects of wind speed. Modified PDTs with wind protective filters have also been developed for deployment in the UK Urban NO_2 Network (UUNN) with an improved measurement accuracy and repeatability. We report the performance of the two designs and also when enclosed in additional shelters. The comparison was carried out against simultaneous reference measurements and was evaluated through a statistical and modelled uncertainty. The model incorporated the individual components of the measurement uncertainty to provide an estimate of the total measurement uncertainty and identified which elements could be reduced across mean values of multiple measurements.

We found that conventional PDTs could be adversely affected by wind speed and that the incorporation of shelters delivered improved repeatability and better accuracy across multiple diffusion tube measurements. The UUNN style diffusive samplers were more accurate than the PDTs and had better repeatability. The additional use of shelters with UUNN style samplers made no discernible difference to the measurements.

1. Introduction

The World Health Organization (WHO) has identified that the combined effects of ambient air pollution and household air pollution are associated with 6.7 million premature deaths a year due to increases in chronic health conditions such as cardiovascular diseases, respiratory diseases, and lung cancer which reduce overall life expectancy (WHO, 2023). The European Union (EU) has implemented clean air policies and set standards to limit the emissions of key pollutants through the implementation of Air Quality Directives (2008/50/EC, UNECE Gothenburg Protocol). Member states and other national governments including the UK (Environment Act, 2021), Norway and Switzerland are closely aligned in their approaches to this issue and have implemented measures to reduce levels of nitrogen dioxide (NO_2) a pollutant originating from the combustion of fossil fuels which is also a key precursor of tropospheric ozone and fine particulate matter, two other key air pollutants that significantly impact public health.

Compliance with legislation is assessed through a combination of

* Corresponding author. National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, United Kingdom.
E-mail address: David.Medland@NPL.co.uk (D. Medland).

https://doi.org/10.1016/j.atmosenv.2025.121219
Received 31 October 2024; Received in revised form 6 March 2025; Accepted 1 April 2025
Available online 2 April 2025

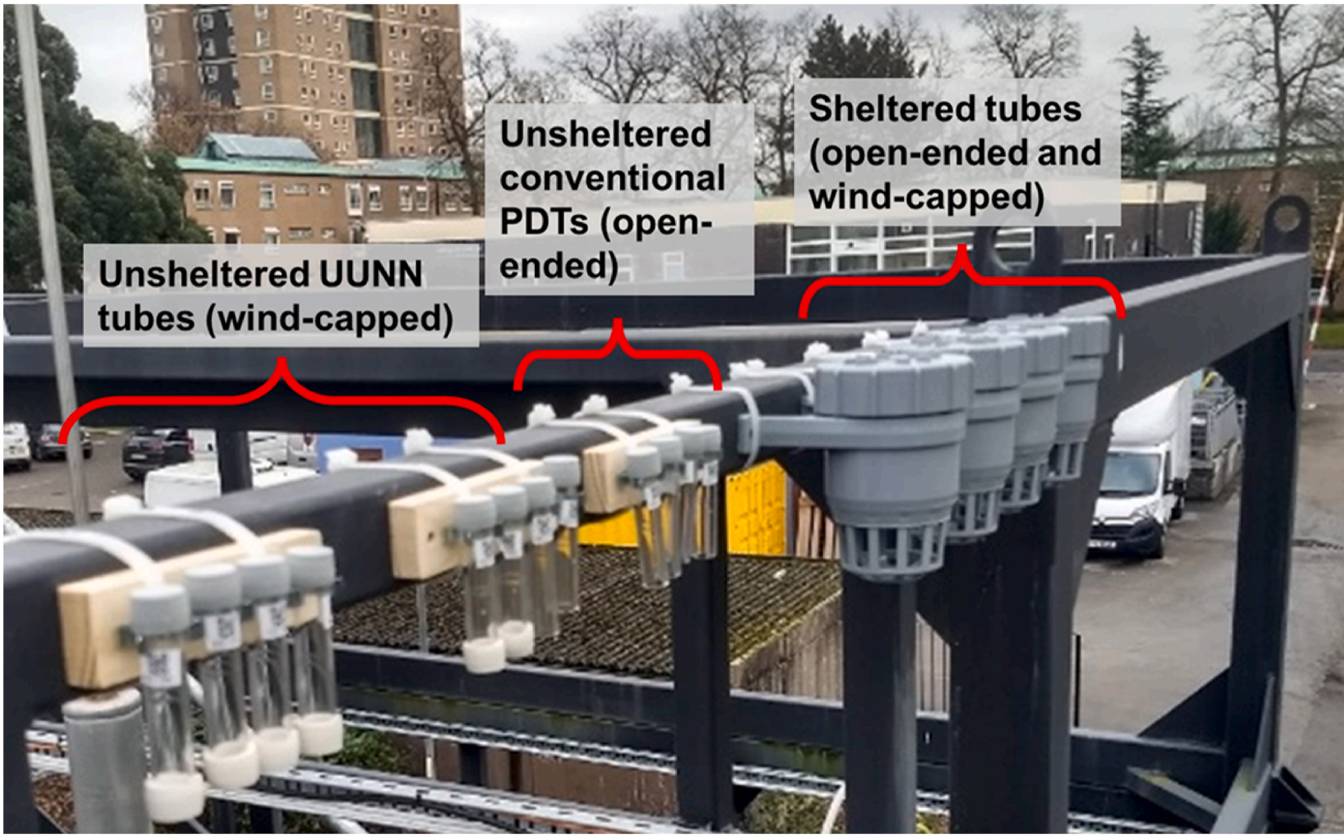

Fig. 1. Unsheltered and sheltered PDT and UUNN tubes deployed at the MAN site.

data from monitoring networks and the results of air pollutant modelling (UK Air, 2025a). For example, in the UK, the Automatic Urban and Rural Network (AURN) deploys reference grade instrumentation to continuously measure the concentration of regulated compounds in over 170 sites (DEFRA, 2022).

One tool used by monitoring networks are Palmes Diffusion Tubes (PDTs) (Palmes et al., 1976). These are cost effective devices widely employed over wide geographical locations for monitoring NO_2, with their low cost providing potential for greater spatial resolution and for citizen science applications (Höhne et al., 2023), However there is some concern on how the method of deployment affects the applicability of the results to determining health impacts, particularly for children (Rowell et al., 2020). They are one of the tools used by UK Local Authorities (LAs) for Local Air Quality Management (LAQM)being used to determine if an area has an annual mean NO_2 concentration below 40 μg m^{-3}. Here, LAs are required to assess air quality (AQ), and declare Air Quality Management Areas where Limit Value exceedances are found and to implement Air Quality Action Plans to deliver improvements. LLAQM, 2019, LAQM, 2022a).

A full technical description of PDTs and their operating principle can be found in the literature (Palmes et al., 1976; Martin et al., 2014) and their use as a monitoring tool is governed by standards (EN 16339, 2013). Briefly, conventional PDTs consist of a sampler body, typically cylindrical, with one end capped and one end left open. The capped end contains a metal grid coated with an adsorbent. When deployed, they are mounted vertically with the open end pointing towards the ground. As ambient air diffuses through the sampler, NO_2 in the air reacts with the adsorbent and is converted to nitrite (NO_2^-), which is later quantified in the laboratory. When deployed conventional PDTs suffer from biases resulting from local conditions and cross interference with other species (Hafkenscheid et al., 2009). These include negative biases from deployment in warm and sunny conditions and relative humidity below a threshold (Heal et al., 2019) and positive biases caused by wind speed, which is the focus of this paper. The current understanding is that wind turbulence at the open end of the PDT causes a shortening of the effective diffusive pathlength of the gas through the tube, resulting in a positive bias in the reported NO_2 (Cape, 2009; Martin et al., 2014). To correct this, in the UK annual bias adjustment factors (specific for each PDT manufacturer) are routinely determined and applied (Butterfield et al., 2021). These factors are calculated annually based on the diffusion tubes collocated with reference instruments and applied to measurements for the following year. Where there are a high number of collocated sites in an area local bias factors may also be used (LLAQM, 2019).

Rather than trying to correct the measurements, wind speed effects can be mitigated by housing samplers in shelters, such as with the Passam sampler, which is based on the PDT and uses specific shelters (Hafkenscheid et al., 2009). When deployed with these shelters the Passam sampler has been found to be suitable for indicative measurements (Rosario et al., 2016). Changes to the PDT design have also been made to reduce the wind effect, Gerboles et al., (2005) found improved performance using PDTs with a Teflon membrane at the open end. More recently modified PDTs with wind protection incorporating an amorphous polyethylene filter at the open end have also been developed and implemented at approximately 300 sites in the UK Urban NO_2 Network (UUNN).

This paper aims to assess the performance of collocated conventional open-ended PDTs (with and without shelters) and PDTs with wind protection (with and without shelters) against reference analysers deployed at various UK AQ monitoring sites. The data acquired was employed to determine the optimum field deployment configurations, and to develop a measurement uncertainty model to identify and quantify the dominant sources of uncertainty.

Table 1
Measurement period of each site with the latitude and longitude used for interpolating reanalysis wind speed data for the uncertainty model US indicates an unsheltered deployment and S a sheltered deployment.

Site	First Measurement	Last Measurement	Deployed monthly	Data Capture	LAT	LON
MAN	Dec 2020	Nov 2021	6 PDT-US 6 PDT-S 6 UUNN-US 6 UUNN S	100 % PDT-US 100 % PST-S 100 % UUNN-US (50 %) 100 % UUNN-S (50 %) Jun–Nov 2021 were found to be erroneous for UUNN >99 % CAPS	53.44294	−2.22081
HOP	Nov 2020	Mar 2022	6 PDT-US 6 PDT-S 6 UUNN-US 6 UUNN-S	82 % PST-US 88 % PDT-S 84 % UUNN-US 87 % UUNN-S No Jan 2021 or Jan 2022 measurements >92 % CAPS	51.44987	−0.04583
MRB	Jan 2021	Dec 2021	3 PDT-US 3 UUNN-US	100 % PDT-US 100 % UUNN-US >93 % Chemiluminescence	51.52269	−0.15557
YOR	Nov 2020	Nov 2021	3 PDT-S 3 UUNN-S	100 % PDT-S 100 % UUNN-S >99 % Chemiluminescence	51.44987	−0.04583

2. Methodology

2.1. Field deployment of diffusion tubes

This study used PDT and UUNN tubes deployed at 4 sites alongside reference instruments - the Honor Oak Park (HOP) London Air Quality Supersite, London Marylebone Road (MRB) AURN site, York Fishergate (YOR) AURN site and at the Manchester (MAN) Air Quality Supersite operated by the University of Manchester, shown in Fig. 1. MAN and HOP are urban background sites while YOR and MRB are urban traffic sites, being around 1–1.5 m away from the road. A description of these site types is available of the UK AIR website (UK AIR, 2025b). All sites reported NO_2 concentrations above the expected minimum detection level of PDTs. Each set of tubes were typically exposed at the sites for 28 days at a height between 2.5 and 4 m before being sent to the manufacturers for analysis. All tubes at the YOR site were deployed inside shelters while at MAN and HOP some tubes in shelters were deployed alongside unsheltered tubes and all tubes at MRB were unsheltered. Table 1 shows which tube types and mounting strategies were used at each location.

.For the period of the tube deployment the MAN and HOP sites used a Teledyne Model T500U Cavity Attenuated Phase Shift NO_2 Analyser. The YOR site used a Teledyne 200 UP chemiluminescence analyser (Diez et al., 2024a) and the MRB site uses a Teledyne API 200E(Alam et al., 2020, EN 14211, 2024). For comparison to the diffusion tube measurements, the reference measurements were converted from ppb to μg m^{-3} where needed and averaged across the period of the diffusion tubes deployment. No gap filling of the reference measurements was done after receiving the data. The reference data for HOP, MAN and YOU are available on Zenodo (Diez et al., 2024b) while the MRB reference data is available on the UK AIR data Archive (https://uk-air.defra.gov.uk/data/)

The National Physical Laboratory provided the required diffusion tubes, mounting blocks and shelters for deployments at Honor Oak Park, Manchester and York. The first 28-day deployment of conventional and UUNN type diffusive samplers (with and without shelters) was carried out in November 2020. The University of York organised the deployments at the roadside site at York and the supersite at Manchester, while NPL was responsible for the deployments at Honor Oak Park and Marylebone Road. The diffusive samplers employed in this study were prepared by Gradko Environmental and analysed using a UKAS accredited colorimetric technique (International Organization for Standardization, 2017; AEA Energy and Environment, 2008; BS EN 13528 1-3:2002/2003) which uses travel and laboratory blanks.

The diffusion tubes measurements used for the comparison with reference instruments were collected between November 2020 and March 2022. Table 1 shows the measurement periods and geographical coordinates for each site. At MAN and HOP 6 tubes of each type were deployed in parallel with and without shelters, while at MRB and YOR only 3 tubes of each type were used.

So that the uncertainty of these measurements could be evaluated and compared, two methods of calculating uncertainty for the annual mean of NO_2 measurements were applied to these data. A statistical method based on guide to the demonstration of equivalence of ambient air monitoring methods (Guide to the Demonstration Report, 2010), which uses the tube measurements and simultaneous reference measurements, and an uncertainty model based on calculating uncertainties for individual elements of the measurement process and combining them into a total measurement uncertainty, with the aim of validating using the statistically calculated uncertainty. This establishes what parts of the measurement process dominate the uncertainty budget and enables the estimation of the measurement uncertainty more widely, e.g. when measurements are combined in longer-term averages. For the comparison of the uncertainty estimates, only measurements taken in 2021 were used.

2.2. Data processing

The concentration measurements from different diffusion tubes at the four sites were divided into four different categories: PDT with traditional mountings (unsheltered), PDT with shelter, UUNN with traditional mountings and UUNN with shelter (see Fig. 1). These sets were compared against the simultaneous reference measurements at each site in order to determine correlation with the reference measurements using gradient and coefficient of determination. These datasets were also compared to each other using an ANalysis Of VAriation (ANOVA) test, which shows if the difference between two datasets is statistically significant. This allows us to determine which tube category performs better by comparing which cases have significant differences and greater agreement with reference measurements. The sheltered and unsheltered UUNN results between June and November 2021 at the MAN site were much higher than the PDT and reference measurements during this time and have been excluded from this analysis and were judged to be erroneous based on an inter-quartile range (IQR) based statistical analysis of the bias of each tube against the reference measurements. Compared to other sites and deployment strategies, where the biases were within bounds, theses were outside the upper threshold of the 3rd IQR+1.5*IQR. Where the UUNN results have been excluded PDT and reference results have also been excluded so that the time periods being compared are equivalent.

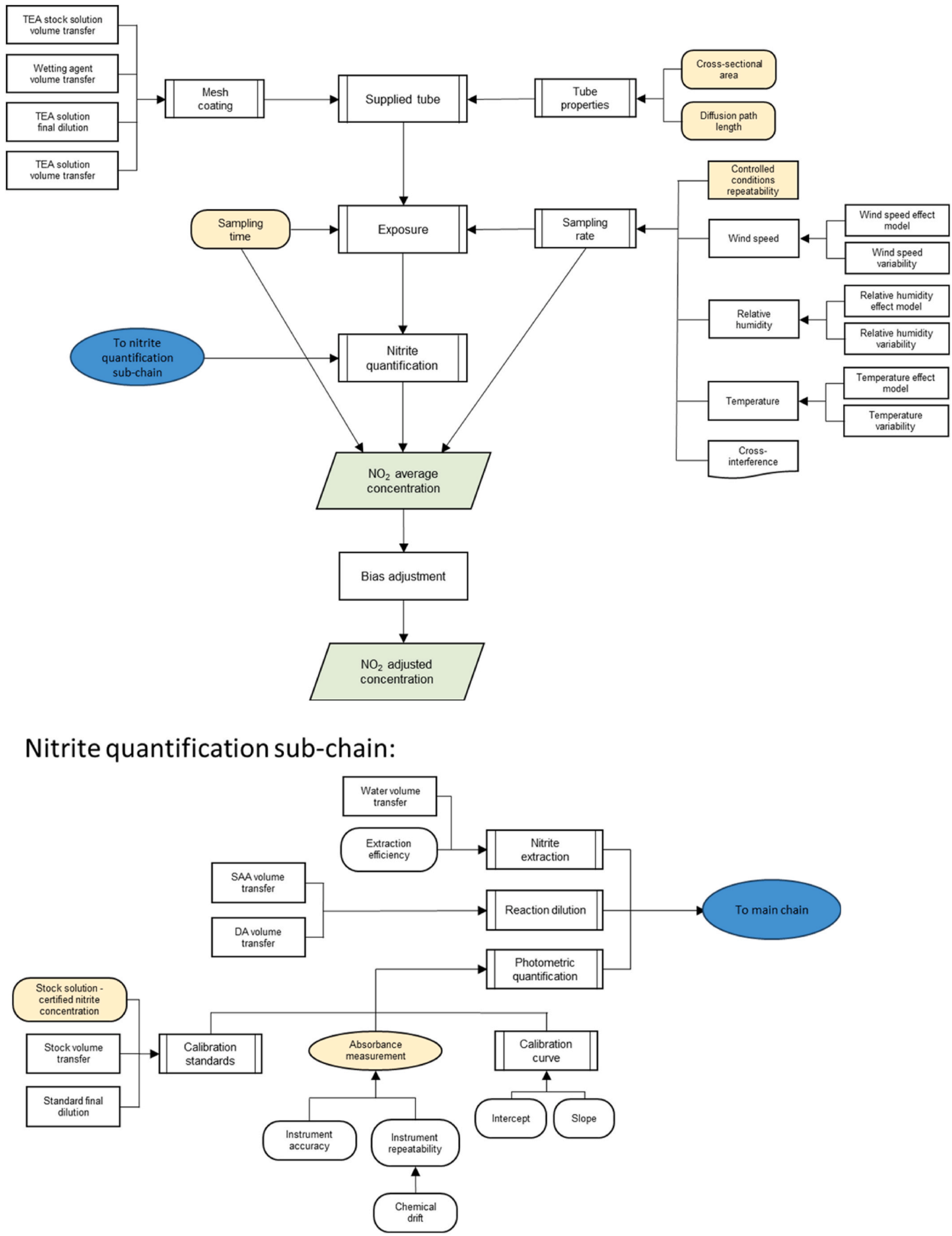

Fig. 2. The traceability chain of individual uncertainty elements for nitrogen dioxide (NO_2) measurements using diffusion tubes from Medland et al. (2025).

2.3. Uncertainty assessment

2.3.1. Statistical uncertainty calculation

This approach is based on the method used to determine agreement between a deployed system, in this case the diffusion tubes, and simultaneous reference measurements. The full method is described in the Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods (Guide to the Demonstration, 2010). This method uses the random uncertainty of the reference system. The manufacturer specification for the CAPS instruments gives a low noise expectation, <0.1 % reading +20 ppt (Teledyne, 2021), while chemiluminescence has a between sampler uncertainty of 15 % at the $k = 2$ coverage level (DEFRA, 2023), but it is assumed to tend towards 0 μg m^{-3} as the random uncertainties are cancelled out in producing a monthly mean (Butterfield et al., 2021).However it should be noted that chemiluminescence instruments are subject to biases from cross-interferences that are not included in this assessment (Alam et al., 2020; Cowan et al., 2024). First, an orthogonal regression is used to find $y_i = a + bx_i$, where y_i is the mean of tube measurements and x_i is the mean of the reference measurement. Then the uncertainty is defined by equation (1):

$$U_{cr}^2(y_i) = \frac{RSS}{(n-2)} + [a + (b-1)x_i]^2 \quad (1)$$

where n is the total number of data points and the residual sum of squares (RSS) is calculated using equation (2):

$$RSS = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \quad (2)$$

The mean of the reference measurements in this case only used reference measurements for months and sites with diffusion tube measurements. The Guidance to Demonstration of Equivalence also includes a method to recalibrate the non-reference method based on the reference measurements, but this was not used in our calculation of uncertainty.

The requirements for LAQM reference measurements are given in the Local Air Quality Management Technical Guidance (TG22) (LAQM 2022b). For NO_2 several instruments that use chemiluminescence and have been approved by the Monitoring CERTification Scheme (MCERTS) are used.

2.3.2. Model uncertainty calculation

The model used was developed using similar principles to the uncertainty calculation for the US Climate reference Network temperature data made available over the Copernicus Climate Change Service Climate Data Store, which are detailed in the data product documentation (Copernicus Climate Change Service, 2023). The steps of the uncertainty calculation are described in a Product Traceability and Uncertainty (PTU) document, produced by first identifying sources of uncertainty and arranged into a traceability chain where the different elements combine up the chain. This method follows similar schemes developed for atmospheric reference measurements and Earth Observation systems (Green et al., 2017). The PTU for NO_2 diffusion tubes used for this model is available in the NPL Report ENV 59 (Medland et al., 2025). Fig. 2 shows the traceability chain for both designs of diffusion tubes used in this study, although the method of calculating uncertainty at certain steps is different depending on the design.

When identifying the different uncertainty elements we considered aspects of the tube design, such as how accurately tube dimensions are known, the deployment method and how local environmental conditions can affect the measurements, the data collection method (in this case, how the nitrite is extracted from the tubes) and data processing (how the extracted nitrite mass is used to produce a final concentration measurement). These individual elements all have their own associated uncertainties found in the literature or experiment.

The individual measurement uncertainties are combined into a monthly uncertainty for each tube and can also be used to calculate site averages, multi-site averages and annual and longer-term averages. To do this the different uncertainty contributions are classified according to how the error is expected to change over different time scales. In this case three error classifications were used as below:

Random – the error changes randomly (asystematically) between every measurement.

Quasi-systematic – the error is systematic over a timescale shorter than that considered by the average, but changes over known timescales.

Systematic – the error does not change or only changes on timescales much longer than the average.

Table 2
Examples of individual uncertainty elements and the classification given to them based on how the error is expected to behave between measurements for the NO_2 diffusion tubes.

Uncertainty element	Site Mean	Multi-Month	Annual	Long term
Cross sectional area	Random	Random	Random	Random
Wind Speed Bias	Systematic	Systematic	Systematic	Systematic
Wind speed Variability	Systematic	Random	Random	Random
Stock solution – certified nitrite concentration	Systematic	Systematic	Quasi-systematic	Random

Table 2 shows the classification of some of the key individual uncertainty elements shown in the traceability chain from Fig. 2. Cross sectional area is a property that varies tube to tube, so it is random between tubes. The wind speed bias is determined from experiments (Martin et al., 2014) and is systematic, whereas the wind speed variability depends on how the wind speed actually varied during the sampling period, so this will be the same for tubes at the same site but be different for each new deployment. The stock solution is used in the calculation of the NO_2 concentration and is systematic while the same solution is used but it becomes quasi-systematic and then random as new batches of solution are made over extended periods. When producing means, random uncertainties and quasi-systematic uncertainties are reduced in magnitude based on the number of measurements or on the number of relevant change periods respectively while systematic uncertainties are not reduced.

For most uncertainty elements the calculation is a relative uncertainty which only requires the diffusion tube NO_2 measurement to calculate. An exception to this is the wind uncertainties in the case of unsheltered PDTs, where the calculation requires the mean wind speed over the sampling period. There was not in-situ wind data readily available for all sites, therefore monthly 10 m wind speeds from the ERA-5 reanalysis (Hersbach et al., 2023) was used for this study. The 4 closest grid points to the site were interpolated linearly to the site location, as shown in Table 1.The monthly mean wind speeds for all the sites was between 1 and 2 m s^{-1}, within the range explored in the laboratory study on which the wind uncertainty calculation was based (Martin et al., 2014). This method allows calculation of uncertainty for single tubes without simultaneous reference measurement but is reliant on appropriate assessment of the individual uncertainty elements.

3. Results and discussion

3.1. Field deployment of diffusion tubes

Fig. 3 shows the concentration of NO_2 from the diffusion tube measurements against the corresponding reference measurements by site, tube and shelter type, while Fig. 4 combines the results from all sites, sorted by tube and shelter type. Linear fits forced through the origin are shown for each data set. This approach is taken because data from blank diffusion tubes and from regular reference monitor calibration indicate that both techniques report 0 in the absence of no NO_2. However, there was some indication in the data of biases at 0 at the London sites.

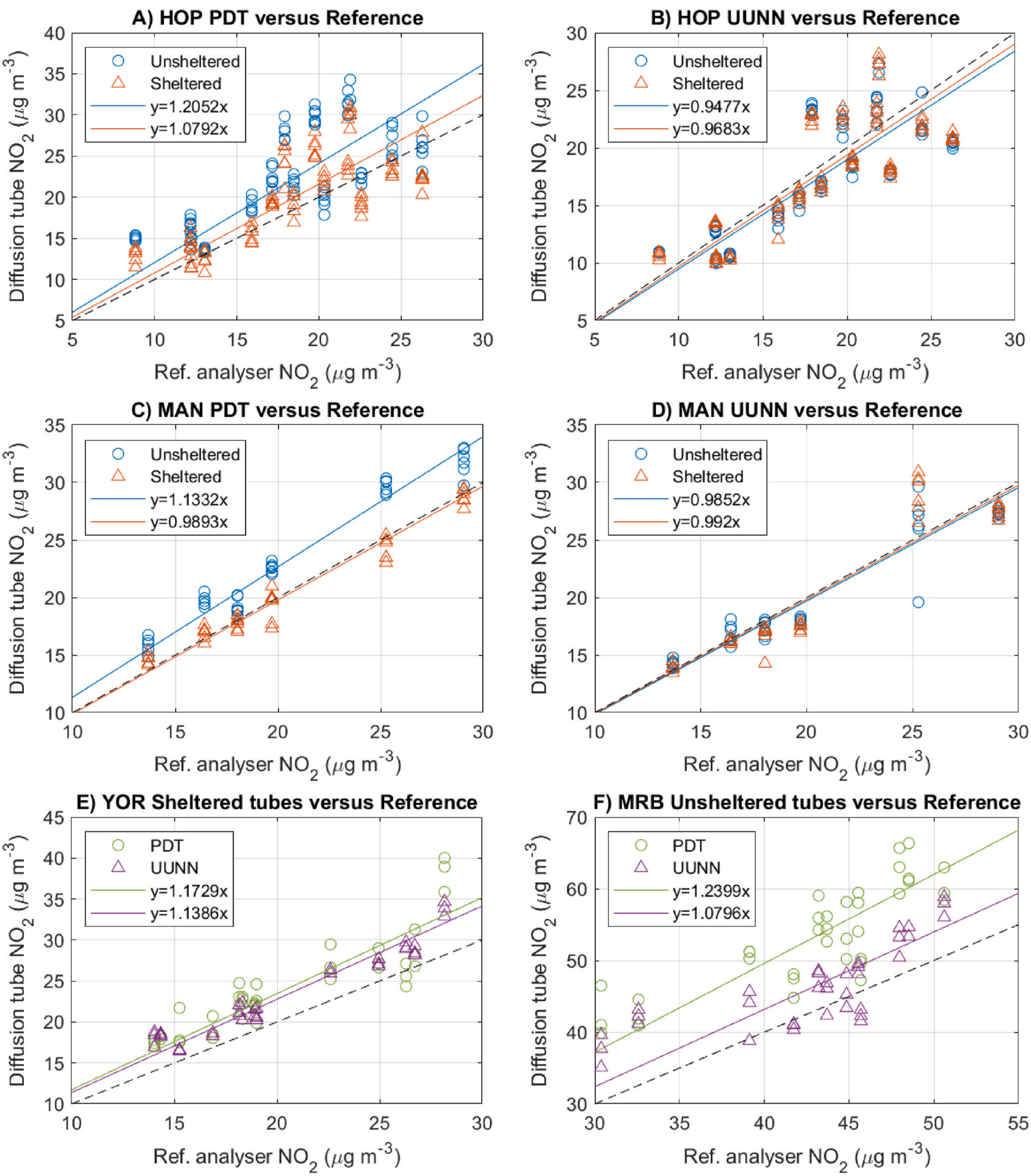

Fig. 3. Correlation plots between tube and reference grade analyser results per site, for PDT tubes at the HOP site(A), UUNN tubes at the HOP site (B), PDT tubes at the MAN site (C), UUNN tubes at the MAN site (D), PDT and UUNN tubes at the YOR site (E) and PDT and UUNN tubes at the MRB site (F). The black dashed line represents the 1:1 line. Lines of best fit are calculated using the least squares method forced through the origin.

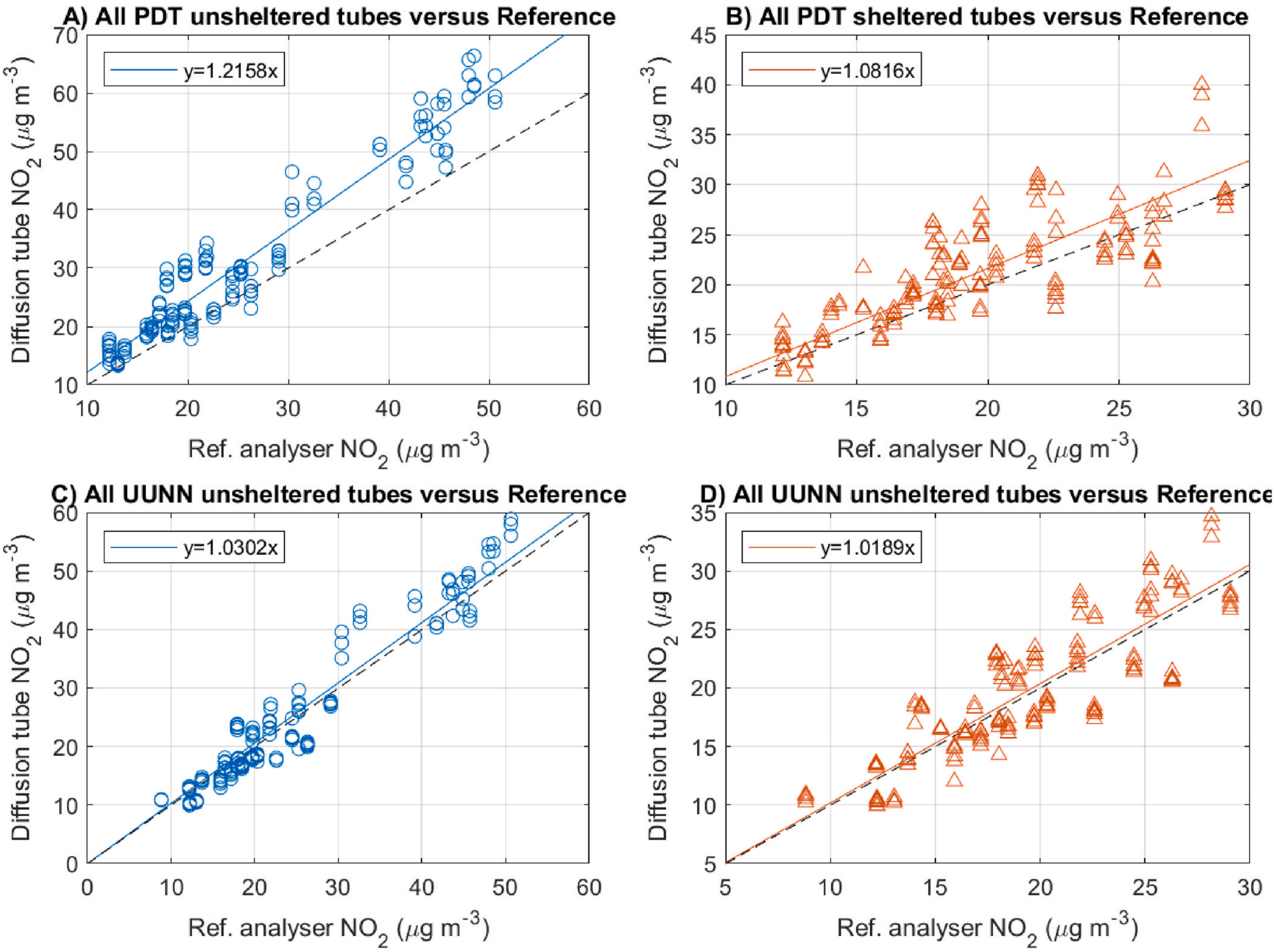

Fig. 4. Correlation graphs between tube and reference grade analyser results for all sites ensembles, for unsheltered PDT tubes (A), sheltered PDT tubes (B), unsheltered UUNN tubes (C) and sheltered UUNN tubes (D). The black dashed line represents the 1:1 line. Lines of best fit are calculated using the least squares method forced through the origin.

Table 3
Results of statistical analysis of significance of different configurations (* indicate a statistically significant effect. US = unsheltered, S = sheltered).

Assessed parameter	Site	Common factor	Compared factors	F calc.	p-value	F crit.
Significance of shelter effect	HOP	PDT	US and S	6.63*	<0.01*	3.90*
		UUNN	US and S	0.41	0.52	3.90
	MAN	PDT	US and S	5.16*	0.03*	3.98*
		UUNN	US and S	0.01	0.92	3.98
Significance of tube design effect	HOP	Unsheltered	PDT and UUNN	34.4*	<0.01*	3.90*
		Sheltered	PDT and UUNN	7.20*	<0.01*	3.89*
	MAN	Unsheltered	PDT and UUNN	5.77*	0.02*	3.98*
		Sheltered	PDT and UUNN	0.01	0.92	3.98
	YOR	Sheltered	PDT and UUNN	0.35	0.56	3.97
	MRB	Unsheltered	PDT and UUNN	18.9*	<0.01*	3.98*
Significance of shelter effect	ALL	PDT	US and S	10.6*	0.00*	3.88*
		UUNN	US and S	1.4	0.23	3.88
Significance of tube design effect	ALL	Unsheltered	PDT and UUNN	11.2*	<0.01*	3.87*
		Sheltered	PDT and UUNN	4.64*	0.03*	3.87*
Equivalence of wind protection	HOP	Site	PDT-S and UUNN-US	11.2*	<0.01*	3.90*
	MAN	Site	PDT-S and UUNN-US	0.1	0.82	3.98
	ALL	Site	PDT-S and UUNN-US	10.56*	<0.01*	3.87*

Compared to the unsheltered PDT tubes, the sheltered PDT and UUNN tubes have gradients closer to unity (within ~10 %). Sheltering the UUNN tubes has a marginal effect on the gradients.

Table 3 shows the ANOVA results for these data sets: the difference between sheltered and unsheltered PDT tubes is considered statistically significant at HOP, MCH and for the combined data from all sites. This, along with the gradients closer to unity in Figs. 3 and 4, indicates a greater agreement between sheltered PDT tubes and reference instruments. No statistically significant difference is found between unsheltered and sheltered UUNN tubes. When comparing between the

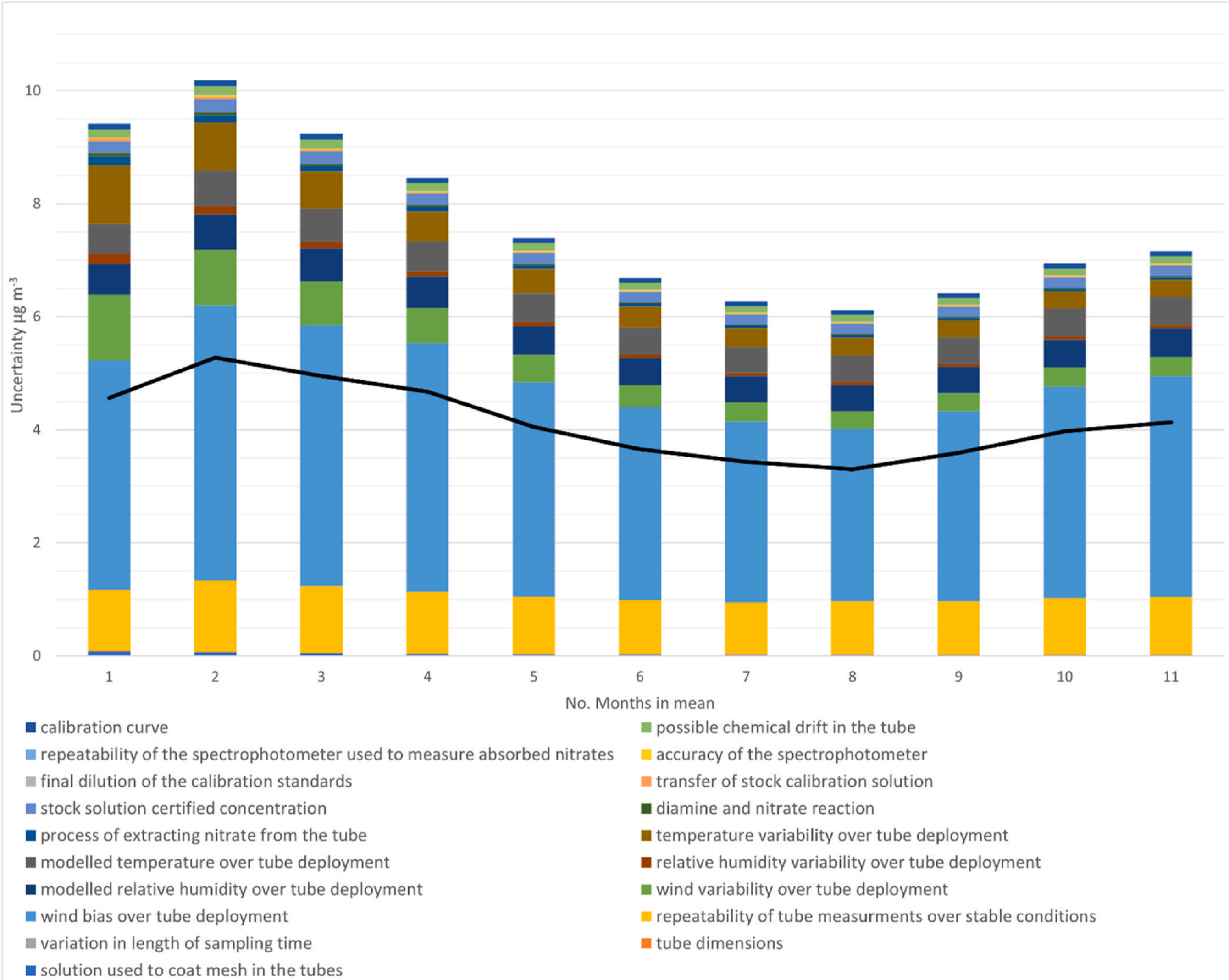

Fig. 5. The value of individual uncertainty elements for 1-month mean to 11-month mean for unsheltered PDT tubes at the HOP site as calculated using the uncertainty model.

different tube types, unsheltered PDT and UUNN tubes show significant differences at all sites and the improved gradient for the UUNN data shows closer agreement with reference measurements. Comparing the sheltered results is more mixed, with statistically significant differences between the two tube designs found in the HOP and the all-site combined datasets, but not in the MAN and YOR datasets. These results indicate improvements in the measurement accuracy from using the UUNN design or mounting PDT tubes in a shelter, however no clear improvement was found when using a shelter for UUNN tubes.

3.2. *Uncertainty assessment*

3.2.1. *Uncertainty from model*

As a result of the different error classifications (Section 2.3 and Table 2), the magnitude of some uncertainty elements is reduced when multiple measurements are aggregated into averages. Therefore, when monthly measurements are combined into an annual mean, the uncertainty is increasingly dominated by the systematic uncertainties as the magnitude of random errors is reduced. This can be seen in Fig. 5, where the different uncertainty elements for the HOP PDT-US dataset are calculated for means from 1 to 11 months. In the total uncertainty calculations these are summed in quadrature, so the combination of the individual uncertainties (each bar stack in Fig. 5) is larger than the total uncertainty (solid black line in Fig. 5). Since the uncertainty elements are calculated from relative uncertainties, the NO_2 concentration affects the shape as there tends to be higher NO_2 concentrations at the beginning and end of the year (*i.e.*, in winter). The largest uncertainty element is the wind bias, the magnitude of which is determined by the wind speed from the reanalysis. This is also the uncertainty that varies the most over the year, between 3.1 and 4.9 µg m^{-3} since it also relies on wind speed. The uncertainty from the variability in wind over time of deployment is the uncertainty that reduces most, from 1.2 to 0.3 µg m^{-3}. Fig. 6 shows the same breakdown of the uncertainty contributions for the UUNN-US tubes at the HOP site. The design of the tube has greatly reduced the wind-related uncertainties and one of the larger elements, the temperature variability from the temperature effect model, is random between different months so it is reduced as more months are included. Starting at 0.7 µg m^{-3} in the one-month mean, the largest single contribution, it then decreases to 0.2 µg m^{-3} in the 11-month mean.

Fig. 7 illustrates the effect of temporal averaging on the overall uncertainty. Here the annual site mean uncertainty for the PDT-US tubes is close to the mean of the site mean and individual tube uncertainties, where for the UUNN-US tubes the annual site mean uncertainty is close to the low end of the shorter-term uncertainties.

One effect that is not apparent from the data shown in Fig. 7 is how

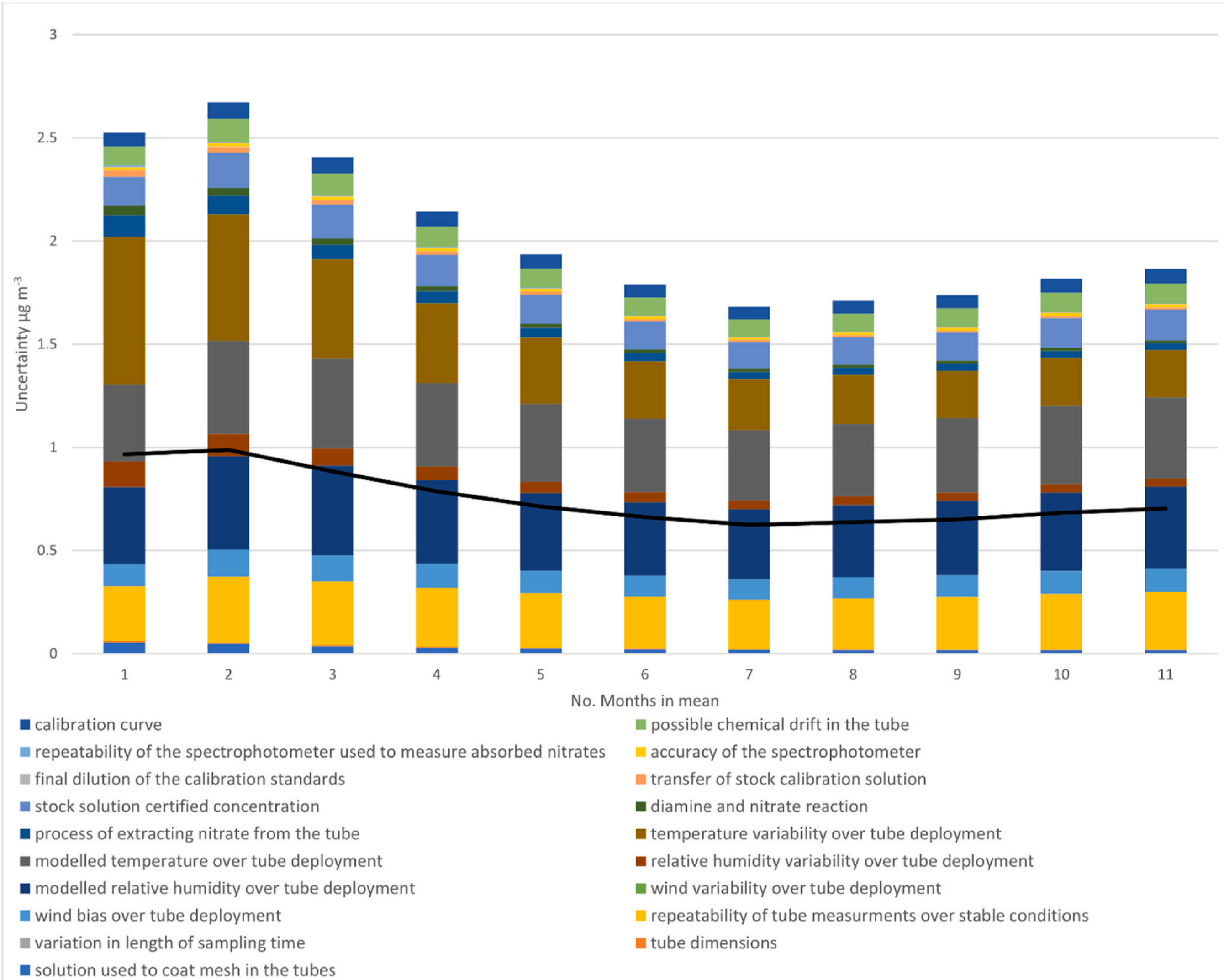

Fig. 6. The value of individual uncertainty elements for 1 month mean to 11 month mean for unsheltered UUNN tubes at the HOP site as calculated using the uncertainty model.

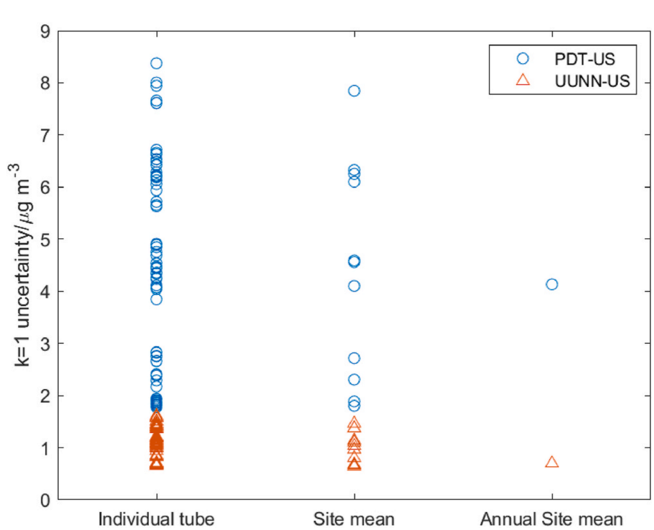

Fig. 7. The individual tube, site mean and annual site mean uncertainties for unsheltered PDT and UUNN tubes at the HOP site calculated using the uncertainty model.

Table 4
The annual mean and annual mean uncertainties calculated using the model and statistical method for all tubes and sites.

Site - tube type - mounting	Tube mean/µg m^{-3}	U(model) $k = 1$/µg m^{-3}	U(stat) $k = 1$/µg m^{-3}	*Diffusion tubes used for mean (N)*
ALL-PDT-US	27.7	5.3	5.8	126
ALL-PDT-S	19.8	1.3	3.2	123
ALL-UUNN-US	24.0	0.9	3.6	125
ALL-UUNN-S	18.9	0.7	3.6	123

the quasi-systematic uncertainties affect the longer-term means. The elements identified as quasi-systematic are only expected to change over 1 or 2 years, when stock solutions used for the tube analysis are re-made or when assumed temperature and relative humidity conditions used for the final value calculation are updated. These produce step reduction in uncertainties according to the number of relevant periods the data is collected over, so if the means were calculated over 2 or 3 years there would be further reductions in the long-term uncertainties that are not seen in the annual mean uncertainty.

Table 5
The annual mean and annual mean uncertainties calculated using the model and statistical method for all tubes and sites with June–November MAN measurements removed.

Site - tube type - mounting	Tube mean/μg m^{-3}	U(model) k = 1/μg m^{-3}	U(stat) k = 1/μg m^{-3}	*Diffusion tubes used for mean (N)*
ALL-PDT-US	30.1	5.7	6.5	96
ALL-PDT-S	20.1	1.3	3.4	93
ALL-UUNN-US	25.2	1.0	3.0	95
ALL-UUNN-S	18.4	0.7	2.6	93

3.2.2. Comparison of uncertainties

Table 4 shows the annual mean NO_2 measurements along with the k = 1 uncertainty calculated from both methods for the all-site mean. There is close agreement in the uncertainties for the unsheltered PDT tubes. The statistical method produces a lower uncertainty for the UUNN and sheltered PDT tubes compared to the unmodified PDT tubes as expected by the model, although the uncertainty calculated from the statistical method is higher than that calculated by the model. Some of this difference may be the result of bias in the reference method not being accounted for in the statistical uncertainty calculations, but more work is needed to understand how much impact this would have.

For the MAN UUNN tubes the unusually high measurements between June and November compared to reference measurements produced a higher uncertainty. Table 5 shows the all site mean uncertainties calculated when these measurements are removed. The UUNN uncertainties are reduced but there is still a much greater reduction in uncertainty between PDT and UUNN tubes in the model uncertainty than the statistical uncertainty. This difference arises from changes in measurement repeatability in stable conditions and wind speed to the sampling rate uncertainty (Figs. 5 and 6). The uncertainties in these cases are based on the laboratory performance of the tubes in the Controlled Atmosphere Testing FACility (CATFAC) (Martin et al., 2014). The conditions repeatability was determined to be 4.8 % for the PDT tubes and 1.7 % for the UUNN tubes based on multiple measurements at constant NO_2 concentration, wind speed, temperature and humidity. The wind speed uncertainties are also based on CATFAC measurements at constant concentration, temperature and humidity but with wind speeds between 0.5 m s^{-1} and 2 m s^{-1}. Based on this, the model used a bias contribution of 2.6 % for PDT-S tubes and 0.6 % for UUNN tubes and no contribution from wind variability for either of these cases. For the PDT-US tubes wind bias was calculated as (2.6 + 9 × wind speed) %, and the variability as (3 × wind speed) %. This method produces quite a large contribution to the uncertainty but since the PDT-US case shows good agreement between the two uncertainty calculation methods, this does not seem unreasonable. The comparison of uncertainty methods suggests that there may be a smaller reduction in the wind effect on sheltered and UUNN tubes in the field than expected from the CATFAC results. It is also possible that other uncertainty elements have been underestimated, or not considered in the analysis, and this becomes more obvious when the influence of the wind effect is reduced. While the unsheltered PDT uncertainty showed smaller differences between the two methods, future work may be able to improve this by using in-situ wind measurements instead of reanalysis, which may be representative of the region but does not account for local effects.

4. Conclusions

Comparison to reference NO_2 measurements shows unsheltered UUNN tubes are more accurate than unsheltered PDT tubes, with the overall measurement bias decreasing from 13-24 % to 1.5–7 %. This improvement was statistically significant individually at all sites, and for the overall combined dataset, indicating that the UUNN tubes have better performance than the PDT tubes. The performance of PDT tubes is greatly improved when these are mounted in a shelter: at the HOP and MAN sites, where unsheltered and sheltered PDT tubes were co-located, the measurement bias was between 0 and 8 % for the sheltered tubes, compared to 13–20 % for the unsheltered tubes. Mounting UUNN tubes in shelters was not found to produce statistically significant improvements in performance compared to UUNN tubes in regular mountings. Comparison between sheltered PDT and UUNN was more mixed, with significant improvement for UUNN over PDT at the HOP site and in the combined dataset but no significant difference in results for the MAN and YOR sites. Unsheltered UUNN tubes had significant improvement over sheltered PDT tubes at the HOP site and in the combined dataset but not at the MAN site. Although not completely consistent, these results indicate there is some advantage to using UUNN tubes over PDT in shelters. These improvements are driven by the tube and shelter design reducing the effect of wind on the measurements.

Uncertainty analysis using statistical methods shows smaller uncertainties for UUNN tubes (sheltered or unsheltered) compared to PDTs. Sheltered PDTs also have smaller uncertainties than unsheltered PDTs. However, the statistical method returned larger uncertainties than the uncertainty model. The model uncertainty was most similar to the statistical uncertainty for the unsheltered PDT tubes (within ~15 %), whereas for sheltered PDTs and UUNN tubes the statistical uncertainty was between 2.5 and 5 times larger than the modelled uncertainty. This may indicate that, while the effect of wind has been significantly reduced, it has not been eliminated for either the UUNN tubes or sheltered PDT tubes, and while the uncertainty from the wind was based on laboratory results, in the field the reduction may not be as great as predicted. Future work could determine these uncertainties using Monte Carlo models, this may enable quasi-systematic to be included and resolve some issues with uncertainty changing with absolute factors such as wind speed (BIPM, 2008).

Overall UUNN tubes are seen to provide better performance compared to reference and lower uncertainty than PDT tubes. Although the PDT tube results can be improved by mounting in a shelter, this improvement is not as much as that provided by the UUNN tubes, which have lower uncertainty and in some cases significantly better agreement with reference measurements.

CRediT authorship contribution statement

David Medland: Writing – original draft, Methodology. **Gabriel Garcia:** Writing – original draft, Methodology. **Tom Gardiner:** Writing – review & editing, Methodology. **Nicholas A. Martin:** Writing – review & editing, Methodology. **Valerio Ferracci:** Writing – review & editing. **Ashley Wilkins:** Data curation. **David Fryer:** Data curation. **Tom Holmes:** Data curation. **Pete Edwards:** Writing – review & editing, Data curation. **Sebastian Diez:** Data curation. **David Butterfield:** Writing – review & editing.

Funding

This work was funded through the Strategic Priority Fund Clean Air Programme.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Hersbach, H. et al., (2018) was downloaded from the Copernicus Climate Change Service (2023).

The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains.

The Manchester Air quality supersite was managed by Dr Michael Flynn, Dr Nicholas Marsden and Dr Thomas Bannan. The London Air Quality Supersite was managed by Dr Max Priestman, Dr Stefan Gillott and Dr David Green. The AURN sites are maintained by Bureau Veritas.

Data availability

The authors do not have permission to share data.

References

AEA Energy and Environment, 2008. Diffusion Tubes for Ambient NO_2 Monitoring: Practical Guidance. AEAT/ENV/R/2504 – issue 1a.

Alam, M.S., Crilley, L.R., Lee, J.D., Kramer, L.J., Pfrang, C., Vázquez-Moreno, M., Ródenas, M., Muñoz, A., Bloss, W.J., 2020. Interference from alkenes in chemiluminescent NOx measurements. Atmos. Meas. Tech. 13, 5977–5991. https://doi.org/10.5194/amt-13-5977-2020.

BIPM, 2008. Evaluation of measurement data — supplement 1 to the "Guide to the expression of uncertainty in measurement" — propagation of distributions using a Monte Carlo method. JGGM 101, 2008.

BS EN 13528, parts 1-3, 2002/2003, Ambient air quality. Diffusive Samplers for the Determination of Concentrations of Gases and Vapours.

Butterfield, D., Martin, N.A., Coppin, G., Fryer, D.E., 2021. Equivalence of UK nitrogen dioxide diffusion tube data to the EU reference method. Atmos. Environ. 262. https://doi.org/10.1016/j.atmosenv.2021.118614.

Cape, J.N., 2009. The use of passive diffusion tubes for measuring concentrations of nitrogen dioxide in air. Crit. Rev. Anal. Chem. 39 (4), 289–310. https://doi.org/10.1080/10408340903001375.

Copernicus Climate Change Service, 2023. Near-surface temperature from US climate reference network since 2006 onward: product user guide and specification (PUGS). https://confluence.ecmwf.int/pages/viewpage.action?pageId=358884472.

Cowan, N., Twig, M.M., Leeson, S.R., Jones, M.R., Harvey, D., Simmons, I., Coyle, M., Kentisbeer, J., Walker, H., Braban, C.F., 2024. Assessing the bias of molybdenum catalytic conversion in the measurement of NO_2 in rural air quality networks. Atmos. Environ. 322. https://doi.org/10.1016/j.atmosenv.2024.120375.

DEFRA, 2022. Air pollution in the UK 2021. https://uk-air.defra.gov.uk/library/annualreport/viewonline?year=2021_issue_1#report_pdf.

DEFRA, 2023. Quality Assurance and Quality Control (QA/QC) Procedures for UK Air Quality Monitoring under the Air Quality Standards Regulations, 2309281140_All_Networks_QAQC_Document_2023.Pdf (defra.gov.uk).

Diez, S., Lacy, S., Urquiza, J., Edwards, P., 2024a. QUANT: a long-term multi-city commercial air sensor dataset for performance evaluation. Scientific Data. https://doi.org/10.1038/s41597-024-03767-2.

Diez, S., Lacy, S., Pete, E., Read, K., Josefina, U., 2024b. QUANT: a three-year, multi-city air quality dataset of commercial air sensors and reference data for performance evaluation. Zenodo. https://doi.org/10.5281/zenodo.10775692.

EN 14211, 2024. Ambient Air. Standard Method for the Measurement of the Concentration of Nitrogen Dioxide and Nitrogen Monoxide by Chemiluminescence.

EN 16339, 2013. Ambient air e Method for the Determination of the Concentration of Nitrogen Dioxide by Diffusive Sampling.

Environment Act 2021, c. 30, https://www.legislation.gov.uk/ukpga/2021/30/contents (accessed 8th April 2025).

Gerboles, M., Buzica, D., Amantini, L., 2005. Modification of the Palmes diffusion tube and semi-empirical modelling of the uptake rate for monitoring nitrogen dioxide. Atmos. Environ. 39 (14), 2579–2592. https://doi.org/10.1016/j.atmosenv.2005.01.012.

Green, P., Gardiner, T., Medland, D., Cimini, D., 2017. GAIA-CLIM guidance note/deliverable D2.6, guide to uncertainty in measurement and its nomenclature. http://www.gaia-clim.eu/sites/www.gaia-clim.eu/files/document/d2_6_final.pdf. July 4th 2024.

Guide to the Demonstration report, 2010. Guide to the Demonstration of Equivalence of Ambient Air Monitoring Methods, Report by an Ec Working Group on Guidance for the Demonstration of Equivalence.

Hafkenscheid, T., Fromage-Mariette, A., Goelen, E., Hangartner, M., Pfeffer, U., Plaisance, H., de Santis, F., Saunders, K., Swaans, W., Tang, Y.S., Targa, J., van Hoek, C., Gerboles, M., 2009. Review of the Application of Diffusive Samplers in the European Union for the Monitoring of Nitrogen Dioxide in Ambient Air, JRC Scientific and Technical Reports, EUR 23793 EN. OPOCE, Luxembourg (Luxembourg), 2009. JRC51106.

Heal, M.R., Laxen, D.P.H., Marner, B.B., 2019. Biases in the measurement of ambient nitrogen dioxide (NO_2) by Palmes diffusion tube: a review of current understanding. Atmosphere 10, 357. https://doi.org/10.3390/atmos10070357.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J.-N., 2023. ERA5 monthly averaged data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.f17050d7, 19-JUN-2023.

Höhne, A., Rieke, S.A.A., Kulicke, M., Huynh, T.-T., Telgmann, M., Frenzel, W., Held, A., 2023. Assessing the spatial distribution of NO_2 and influencing factors in urban areas-passive sampling in a citizen science project in Berlin Germany. Atmosphere 14, 260. https://doi.org/10.3390/atmos14020360.

International Organization for standardization, 2017. General Requirements for the Competence of Testing and Calibration Laboratories. ISO/IEC 17025:2017.

Local Air Quality Management - Policy Guidance (PG22), 2022. Defra and the devolved administrations. https://laqm.defra.gov.uk/wp-content/uploads/2023/11/LAQM-Policy-Guidance-2022.pdf.

Local Air Quality Management - Technical Guidance (TG22), 2022. Defra and the devolved administrations. https://laqm.defra.gov.uk/wp-content/uploads/2022/08/LAQM-TG22-August-22-v1.0.pdf.

London Local Air Quality Management , 2019, Technical Guidance (LLAQM.TG(19)) https://www.london.gov.uk/sites/default/files/llaqm_technical_guidance_2019.pdf.

Martin, N.A., Helmore, J., White, S., Snook, I., Parish, A., Gates, L., 2014. Measurement of nitrogen dioxide diffusive sampling rates for Palmes diffusion tubes using a controlled atmosphere test facility (CATFAC). Atmos. Environ. 94, 529–537. https://doi.org/10.1016/j.atmosenv.2014.05.064.

Medland, D., Garcia, G., Martin, N., Gardiner, T., 2025. Methodology for uncertainty reporting in Air Quality studies. NPL Report ENV 59. https://doi.org/10.47120/npl.ENV59.

Palmes, E.D., Gunnison, A.F., DiMattio, J., Tomczyk, C., 1976. Personal sampler for nitrogen dioxide. Am. Ind. Hyg. Assoc. J. 37 (10), 570–577. https://doi.org/10.1080/000288976850752.

Rosario, L., Pietro, M., Francesco, S.P., 2016. Comparative analyses of urban air quality monitoring systems passive sampling and continuus monitoring stations. Energy Proc. 101, 321–328.

Rowell, A., Terry, M.E., Deary, M.E., 2020. Comparison of diffusion tube-measured nitrogen dioxide concentrations at child and adult breathing heights: who are we monitoring for? Air Qual. Atmos. Health 14, 27–36. https://doi.org/10.1007/s11869-020-00909-4.

Teledyne, 2021. T500U specifications. https://www.teledyne-api.com/en-us/Products_/Documents/Specification/SAL000078J%20-%20T500U.pdf. February 27th 2025.

UK AIR, 2025a. Air modelling for defra. https://uk-air.defra.gov.uk/research/air-quality-modelling?view=modelling. February 26th 2025.

UK AIR, 2025b. Site environment types. https://uk-air.defra.gov.uk/networks/site-types. February 26th 2025.

WHO, 2023. Household Air pollution fact sheets. https://www.who.int/news-room/fact-sheets/detail/household-air-pollution-and-health. July 4th 2024.