



## The air quality impacts of pre-operational hydraulic fracturing activities

Shona E. Wilde<sup>a,\*</sup>, James R. Hopkins<sup>b</sup>, Alastair C. Lewis<sup>b</sup>, Rachel E. Dunmore<sup>a</sup>, Grant Allen<sup>c</sup>, Joseph R. Pitt<sup>c,1</sup>, Robert S. Ward<sup>d</sup>, Ruth M. Purvis<sup>b</sup>

<sup>a</sup> Wolfson Atmospheric Chemistry Laboratories, University of York, York YO10 5DD, UK

<sup>b</sup> National Centre for Atmospheric Science, University of York, York YO10 5DD, UK

<sup>c</sup> Department of Earth and Environmental Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK

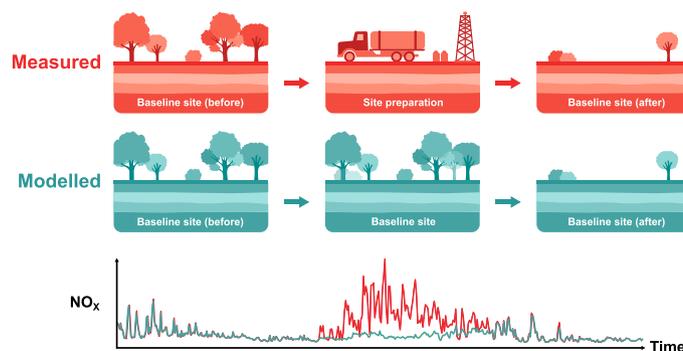
<sup>d</sup> British Geological Survey, Environmental Science Centre, Nicker Hill, Keyworth, Nottingham NG12 5GG, UK



### HIGHLIGHTS

- Prior to hydraulic fracturing there is an extensive period of preparation.
- Pre-operational activities led to an increase in NO<sub>x</sub> (274 %) and decrease in local O<sub>3</sub> (29%).
- Combustion-related sources are responsible for higher primary NO<sub>2</sub> emissions, which may exceed WHO guidelines.
- The pre-operational phase should be included in environmental assessments of shale gas extraction.

### GRAPHICAL ABSTRACT



### ARTICLE INFO

Editor: Hai Guo

#### Keywords:

Fracking  
Oil and gas emissions  
NO<sub>x</sub>  
Baseline monitoring  
Machine learning  
Random Forest

### ABSTRACT

Hydraulic fracturing (fracking) is a short phase in unconventional oil and natural gas (O&G) development. Before fracking there is a lengthy period of preparation, which can represent a significant proportion of the well lifecycle. Extensive infrastructure is delivered onto site, leading to increased volumes of heavy traffic, energy generation and construction work on site. Termed the “pre-operational” period, this is rarely investigated as air quality evaluations typically focus on the extraction phase. In this work we quantify the change in air pollution during pre-operational activities at a shale gas exploration site near Kirby Misperton, North Yorkshire, England. Baseline air quality measurements were made two years prior to any shale gas activity and were used as a training dataset for random forest (RF) machine learning models. The models allowed for a comparison between observed air quality during the pre-operational phase and a counterfactual business as usual (BAU) prediction. During the pre-operational phase a significant deviation from the BAU scenario was observed. This was characterised by significant enhancements in NO<sub>x</sub> and a concurrent reduction in O<sub>3</sub>, caused by extensive additional vehicle movements and the presence of combustion sources such as generators on the well pad. During the pre-operational period NO<sub>x</sub> increased by 274 % and O<sub>3</sub> decreased by 29 % when compared to BAU model values. There was also an increase in primary emissions of NO<sub>2</sub> during the pre-operational phase which may have implications for the attainment of ambient air quality standards in the local surroundings. Unconventional O&G development remains under discussion as a potential option for improving the security of supply of domestic energy, tensioned however against significant environmental impacts. Here we demonstrate that the preparative work needed to begin fracking elevates air pollution in its own right, a further potential disbenefit that should be considered.

\* Corresponding author.

E-mail address: [shona.wilde@york.ac.uk](mailto:shona.wilde@york.ac.uk) (S.E. Wilde).

<sup>1</sup> Now at: School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK.

<http://dx.doi.org/10.1016/j.scitotenv.2022.159702>

Received 15 August 2022; Received in revised form 4 October 2022; Accepted 21 October 2022

Available online 27 October 2022

0048-9697/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Hydraulic fracturing or “fracking” is a technique used in unconventional oil and natural gas (O&G) development. Fracking is the industrial process of hydrocarbon extraction from shale rock formations by injecting large quantities of fluid at high pressure down a well, causing the rock to fracture and thus enabling the flow of trapped gas (Staddon et al., 2016). A combination of technological breakthroughs, such as horizontal drilling, has led to a wide-scale uptake of this technique, since it facilitates the extraction of O&G trapped within shale that cannot be exploited through conventional methods (Archibald et al., 2018). Shale gas has become a key source of natural gas in the United States (US) since 2000, accounting for 75 % of total US dry natural gas production in 2019 (U.S. Energy Information Administration, 2020). Subsequently, interest has spread to other countries, including Australia, Germany and the UK, as this technique has the potential to alter the energy landscape of a country and potentially enhance domestic energy security.

Significant environmental concerns associated with the impact of fracking have accompanied the increase in commercial application. These are most directly centred around sub-surface issues, such as the potential to cause earthquakes (Ellsworth, 2013; Mediaview, 2012) and the possible contamination of water supplies (Vengosh et al., 2014). Wider concerns around the climate impact of combustion of the fracked hydrocarbon are also highly relevant, and also the leakage of methane (CH<sub>4</sub>) during extraction (Zhang et al., 2020; Alvarez et al., 2012, 2018). Unconventional O&G development has also previously been found to have an impact on local and regional air quality. The predominant component of natural gas is CH<sub>4</sub>, a greenhouse gas with a high global warming potential, meaning it is often the focus for climate mitigation policies surrounding the O&G industry (Boucher et al., 2009). CH<sub>4</sub> emissions also contribute to poor air quality through the generation of ozone (O<sub>3</sub>), a secondary pollutant with adverse health effects (Lippmann, 1989; Zhang et al., 2019). In addition to CH<sub>4</sub>, unconventional O&G development leads to emissions of nitrogen oxides (NO<sub>x</sub> = NO + NO<sub>2</sub>) and non-methane volatile organic compounds (VOCs), resulting from point source, mobile and fugitive emissions (Field et al., 2014). Emissions of NO<sub>x</sub> are predominantly linked to the numerous emission sources associated with combustion (Vinciguerra et al., 2015). These include engines from drilling rigs, compressors, and generators, in addition to heaters and pumps. Acute exposure to NO<sub>2</sub> has been widely linked to adverse health effects such as reduced lung function and increased risk of stroke (Shah et al., 2015). Furthermore, both NO<sub>x</sub> and VOCs are recognised as key pollutants in the production of O<sub>3</sub>. Consequently, elevated levels of O<sub>3</sub> in the atmospheric surface layer have been linked to emissions from regions of O&G production (Edwards et al., 2014; Helmig et al., 2014).

Often, the focus on air quality emissions relates to releases associated with the opening of a well and the subsequent extraction of gas. Previous work conducted in the US identified drilling and flaring to be the dominant sources of NO<sub>x</sub> (Dix et al., 2020). However, prior to drilling and extraction there is a significant period of preparation, during which the well pad must be built, the rig transported and constructed and material required for fracking must be transferred onto site. This results in a considerable increase in heavy duty vehicle traffic. The environmental impacts of road traffic emissions associated with unconventional O&G operations are often noted but rarely quantified (King, 2012). Previous modelling work has shown that traffic related to hydraulic fracturing could lead to 18 %–30 % increases in total daily NO<sub>x</sub> emissions (Goodman et al., 2016). Moreover, it is apparent that the enhancement above baseline values is most significant for rural or village locations, where concentrations are typically low, but these are where future shale gas developments would likely occur. The future for shale gas exploitation in the UK remains uncertain, although models have been used to estimate impacts in the absence of any commercial scale activity. One modelling study demonstrated that increases of NO<sub>x</sub> and VOC emissions associated with hydrocarbon extraction could lead to approximately 110 extra premature deaths a year in the UK from increases of up to 30 ppb in the monthly mean of daily 1-hour maximum NO<sub>2</sub>

(Archibald et al., 2018). Understanding the totality of emissions from green field site through to well completion are vital to help inform decision making and future policies, should the sector be considered for expansion (Purvis et al., 2019).

### 1.1. UK context

Permits for shale gas extraction in the UK are currently paused and have gone no further than early exploratory stages. As of November 2019, the UK government announced a moratorium on fracking in England due to unpredictable seismic events (Priestley, 2020). However, increasing energy costs as a result of the Ukraine conflict in early 2022 have provoked renewed interest in the industry (BBC News, 2022c). In April 2022 a UK scientific review into shale gas extraction was launched to assess the progress made to address safety concerns surrounding the industry (The Guardian, 2022). In an effort to decrease reliance on imported energy, the UK government overturned the ban in September 2022, stating that fracking would be allowed in areas where there is local community support (BBC News, 2022a, 2022b).

To date, exploration has taken place at a small number of sites in the UK. These include in Kirby Misperton (KM), North Yorkshire, where there has been extraction from a conventional gas field, named the KM1 well-site, for more than decade. In 2013, an extension to KM1 was constructed and a new well was drilled, referred to as KM8, where approval was initially granted for fracking in May 2016.

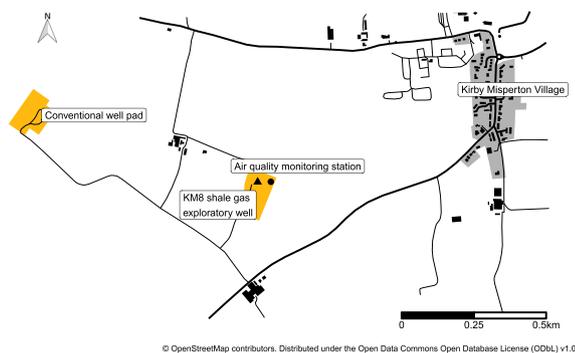
Following approval, significant changes in the on-site infrastructure occurred at KM during September 2017. Machinery required for hydraulic fracturing was brought onto the well pad in preparation for the start of operational activities. Drilling rigs, pumps, compressors, diesel generators and containers holding water, sand and fracking fluid were among the equipment transported onto the site. In addition to the increase in equipment and activity on the site itself, traffic volume due to delivery trucks increased along with additional idling vehicles in close vicinity to the site from protest activities as well as a high volume of policing and media interest. This phase of preparation is defined from here on as the “pre-operational” period. Despite the preparations, final government consent was never received and all fracking-related operations subsequently ceased at KM in February 2018 (BBC News, 2019). The isolated nature of the pre-operational period presents a unique opportunity to assess a relatively understudied stage of the well pad life-cycle. Currently, there are only a small number of studies which report on emissions during the pre-production stages of O&G development (Hecobian et al., 2019; Jarosławski et al., 2022). Moreover, fracking is a temporary process and usually takes only 3–5 days for a single well once drilled. This is short in comparison to the length of the preparation period, which takes place over a number of weeks or months. In the U.S., where large multi-well pads are common, numerous wells (10–20) can be fracked on timescales comparable to the preparation period. In both cases, pre-operational emissions would be expected to occur for a significant proportion of the entire extraction process and are thus important to consider.

### 1.2. Objectives

The primary objective of this study is to provide a quantitative reference for the impact of pre-operational hydraulic fracturing activities on local air quality. We apply statistical predictive models in a unique context to quantify the associated change in air pollutant mixing ratios (NO<sub>x</sub> and O<sub>3</sub>) during site preparation, whilst controlling for local meteorological conditions.

## 2. Data and methods

An air quality monitoring station was installed along the east wall of the well site, approximately 45 m from the KM8 well head, shown in Fig. 1. The enclosure was positioned to be predominantly downwind of the shale gas extraction infrastructure, whilst being open and unobstructed in all wind directions. The close proximity of the monitoring station to the well head



**Fig. 1.** Locations of the baseline monitoring station (circle) and the KM8 well (triangle). Lines identify major and minor roads in the area. Grey shading shows residential areas and the gold shading shows well pads operated by Third Energy.

provided a high sensitivity of observations to operational activity. Instrumentation was housed in a mains-powered, air-conditioned, weather-proof enclosure. Ambient air was sampled from gas phase inlets which were fixed to the top of the monitoring station at a height of 3 m.

### 2.1. Instrumentation

The monitoring station was equipped with a suite of air quality instrumentation, allowing the measurement of several air pollutants along with meteorological variables. Data was collected at a resolution of 1-minute but was hourly aggregated for use in this analysis. A summary of the instrumentation is provided in Table 1.

Quality assurance (QA) and quality control (QC) procedures were routinely performed for all aspects of data acquisition, including; equipment evaluation, site operation and maintenance and data review. Calibrations of air quality instrumentation were conducted on a monthly basis throughout the entirety of the measurement period. All gas phase instrument calibrations were traceable through a chain to international reference standards to maintain a high accuracy and provide known uncertainties in the recorded data. This also ensures comparability with similarly calibrated instrumentation, such as those parts of the UK's monitoring networks.

On-site span and zero point calibrations were performed monthly for the  $\text{NO}_x$  analyser. The span calibration was conducted using a 100 ppb  $\text{NO}$  standard in  $\text{N}_2$ , linked to a National Physical Laboratory (NPL) binary standard and also referenced to the WMO Global Atmospheric Watch (GAW) scale. Zero calibrations were performed using an air scrubber filled with Sofnofil followed by activated charcoal. The  $\text{NO}_2$  conversion efficiency was calculated on an annual basis by returning the instrument to the laboratory to carry out a gas phase titration with known quantities of  $\text{O}_3$ .

The  $\text{O}_3$  instrument provides an absolute measurement but was verified annually off-site using a Model 49i-PS Primary Standard over the calibration range 0–500 ppb. The primary standard was itself checked annually against a certified source by NPL. Instrument blanks took place monthly using air filtered through an activated charcoal trap. Instrument blanks

**Table 1**  
Instrumentation details for the air quality monitoring station at KM.

Species	Instrument	Measurement technique	Frequency	Precision
Air temperature, air pressure, relative humidity (RH)	Luft WS500-UMB compact weather station	Temp: PT100 measuring element, Pressure: resonant pressure transducer, RH: heated capacitive sensor element	1 min	Temp = $\pm 0.1$ C, Pressure = $\pm 0.1$ hPa, RH = $\pm 2$ %
Wind speed, wind direction	Gill 2D WindSonic	Anemometer	1 s	Speed = $\pm 2$ % at $12 \text{ m s}^{-1}$ , Direction = $\pm 2$ at $12 \text{ m s}^{-1}$
$\text{NO}$ , $\text{NO}_2$	Teledyne T200UP	Chemiluminescence with photolytic converter	1 min	$\text{NO} = 0.65$ ppb, $\text{NO}_2 = 0.75$ ppb
$\text{O}_3$	ThermoFisher Model 49i Ozone Analyser	UV absorption	1 min	0.74 ppb

generally read between  $-0.5$ – $0.5$  ppb, resulting in a maximum uncertainty of 7 % for typical daytime mixing ratios of  $\text{O}_3$  (15–30 ppb).

### 2.2. Random forest models

When considering changes in ambient air pollution, it is often difficult to disaggregate changes in mixing ratios due to meteorology from a change in the number or strength of emission sources. Baseline data collected prior to the period of interest can be exploited to identify events that deviate from the “normal” (Shaw et al., 2019), however the influence of meteorology often adds complications, making the quantification of such events challenging. Controlling for meteorological variability allows deviation events to be more robustly assessed. This is achieved by training a statistical model where a range of explanatory variables can be used to account for some of the variability in pollutant mixing ratios.

Random forest (RF) models have been widely used elsewhere to control for the effects of weather in air quality datasets, predominantly in the application of a “meteorological normalisation” technique (Grange et al., 2018; Grange and Carslaw, 2019; Zheng et al., 2020; Cole et al., 2020).

The method here is somewhat different since the models are used to predict mixing ratios during the pre-operational period, assuming a business as usual (BAU) scenario. This is essentially an intervention study, similar to other work quantifying the effect of an airport closure (Carslaw et al., 2012) and more recently the effect of the COVID-19 lockdowns on air quality (Forster et al., 2020; Grange et al., 2021; Carslaw, 2020). The BAU scenario assumes pre-operational activities did not occur at the site and therefore baseline conditions were uninterrupted and continuous. The BAU scenario is then compared with observations to quantify the incremental effect of the pre-operational period on air quality at KM.

#### 2.2.1. Model construction

RF models were developed for  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{NO}_x$ ,  $\text{O}_3$  and total oxidant ( $\text{OX} = \text{NO}_2 + \text{O}_3$ ) using the **rmweather** R package (R Core Team, 2020; Grange et al., 2018). Models were trained using hourly-averaged baseline data collected before and after the pre-operational period, which was an isolated period of activity on site (Fig. A.1). Of this training set, 80 % of the input data was used for model training whilst the remaining 20 % was used for model validation. For each species this split equated to approximately 19,000 training and 4800 testing observations. The performance of such models was assessed before they were used to predict pollutant mixing ratios using local meteorological variables as the model input. The model parameters were set as follows: the number of trees was fixed at 300, the minimum node size was set to 5 and the number of independent variables randomly sampled at each split was 3 (the square route of the number of independent variables). The explanatory variables used for prediction were: Unix date (number of seconds since 1970-01-01) as the trend term, Julian day as the seasonal term, weekday, hour of day, air temperature, atmospheric pressure, wind direction and wind speed. For the input meteorological variables, missing data was replaced with the median. An additional variable, “section”, was introduced with acted as an identifier for data “before” and “after” the pre-operational period. This variable essentially helps account for the fact that the baseline characteristics after the pre-operational period may not be identical to those before. For example,

additional security remained on site after pre-operational activities ceased. Therefore this variable provides the models with a way of distinguishing between the two sections of data when making predictions during the baseline period. However, the “section” variable was excluded when predicting mixing ratios during the pre-operational period since it would take neither value that was used in the training process.

### 3. Results and discussion

#### 3.1. Model performance

The RF models performed well, with  $R^2$  values ranging from 60 % to 90 % (Table B.1). This suggests that the variation in mixing ratios of these pollutants can be reasonably well explained by a combination of meteorological conditions, along with time variables, which essentially act as proxies for emission source strength (Derwent et al., 1995).

The performance of the RF models was validated using a set of baseline data which was held back from the training process. Before initialising the models, the baseline data was randomly split into “training” and “testing” sets of data, accounting for 80 % and 20 % of observations respectively. Since the testing set of data was not used to build the models it can be used to provide insight into how well the models generalise to an independent data set. The models performed well with  $R^2$  values ranging from 58 % to 90 %, suggesting the models are suitably capable of predicting unseen data (Fig. A.2). The best performing RF model was for  $O_3$ , which had an  $R^2$  matching that of the training set (0.9) and for which data are closely scattered around the 1:1 line (Fig. A.2). Poorer model performance was found for  $NO_x$ , in particular NO. Data below a measured value of 20 ppb is well correlated around the 1:1 line but the model fails to predict short-lived spikes in mixing ratios (Fig. A.2). This is perhaps to be expected

since NO is a fast-reacting primary pollutant, where enhancements are strongly linked to events in the local environment in the vicinity of the monitoring site, such as a passing vehicle. Proxies in the model such as hour or day attempt to control for this but are unlikely to be good predictors for sporadic events, hence model performance is expected to be weaker. A much better performance was seen for  $NO_2$  since this is predominantly a secondary pollutant.  $NO_2$  mixing ratios are driven by air originating from more widespread sources on larger spatial scales, such as local traffic flow, which the proxies in the model capture much better.

The RF models were further evaluated by looking at the relative importance or predictive power of each independent variable. This metric is calculated by first assessing the model performance by passing a validation set of out of bag (OOB) data through the trained model. The model accuracy is then computed by comparing the predicted values to the observed values in the validation data set. Next, the values contained within the column of a single variable are permuted or randomly shuffled, essentially giving them no predictive power. The validation data are then passed through the RF model again and the performance evaluated. The feature importance is essentially the decrease in prediction accuracy caused by permuting the column (Breiman, 2001). The importance of each variable is averaged across all trees to obtain the permutation importance for the entire forest (Strobl et al., 2008). Fig. 2 shows the permutation importance of each predictive variable for each pollutant. The trend term (unix time) and seasonal term (Julian day) were the most important explanatory variables for both components of  $NO_x$ , suggesting  $NO_x$  concentrations at KM are largely driven by annual cycles in regional emissions. Interestingly, hour-of-day and day-of-week were found to have little influence on the models ability to predict  $NO_x$ , suggesting time variables are relatively weak proxies for local emission source strengths, such as traffic, in a rural location such as KM. Similarly, wind direction was a relatively unimportant variable,

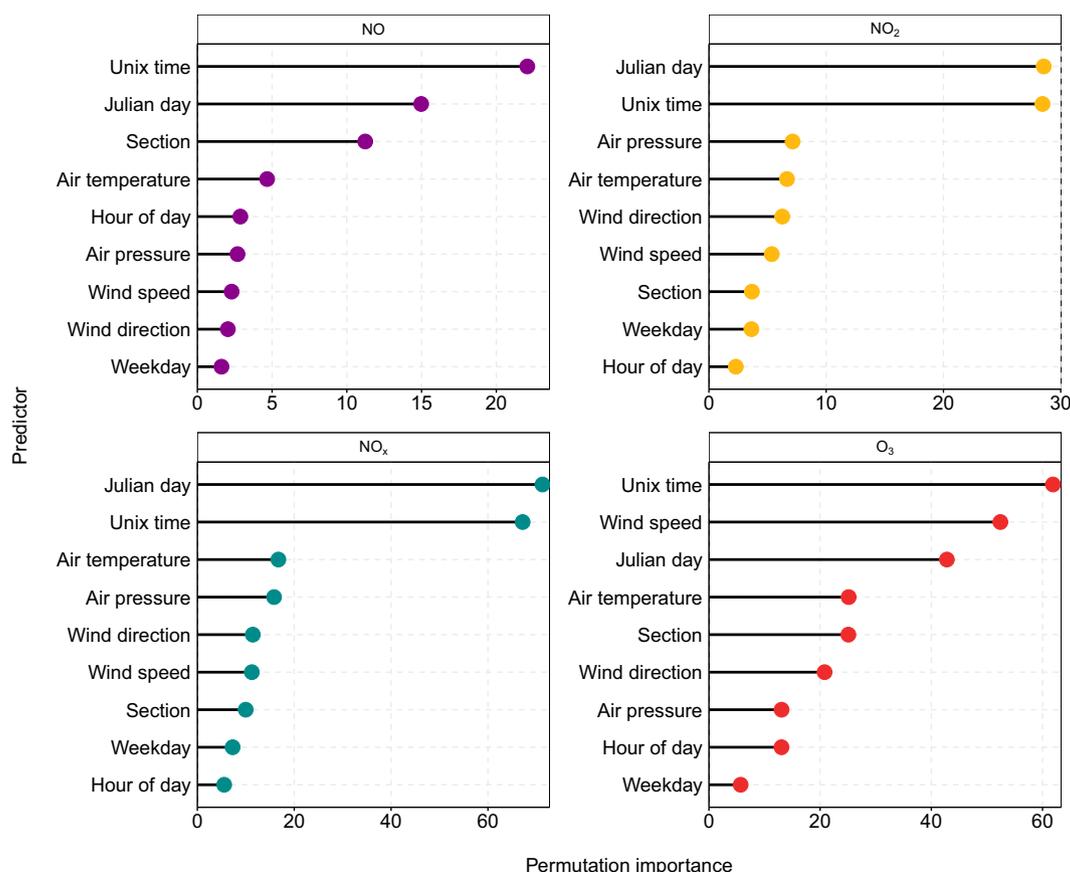


Fig. 2. Variable importance plot for 300 RF models for NO, NO<sub>2</sub>, NO<sub>x</sub> and O<sub>3</sub> at KM.

again reflecting the characteristics of a rural background site where concentrations are not influenced by specific point sources of emissions but rather by the integrated contribution from all upwind sources. In terms of  $O_3$ , wind speed was the second most important variable, which is consistent with a polar plot of  $O_3$  (Fig. A.3c), where high  $O_3$  is associated with relatively strong wind speeds ( $8\text{--}10\text{ m s}^{-1}$ ) from the west.

### 3.2. Effect of the pre-operational period on ambient mixing ratios

The impact of the pre-operational period on  $NO_x$  mixing ratios was initially investigated by studying the variation in pollutant mixing ratios by each hour of the day. Pollutant concentrations are often influenced by the structure and diurnal variability of the planetary boundary layer (PBL). Surface heating drives the formation of the PBL. As the PBL grows throughout the day, pollutants are diluted as they mix with cleaner air from the free troposphere. Similarly, as the PBL shrinks at night time in the absence of surface heating, emissions become more concentrated as they are confined to a smaller volume of the atmosphere. Fig. 3 shows the average diurnal mixing ratios of NO,  $NO_2$  and  $NO_x$  throughout the baseline and pre-operational periods. The baseline data for each year was filtered to the equivalent of the pre-operational period (19th September–1st February) in order to prevent bias due to the seasonal variation of  $NO_x$  mixing ratios. Additionally, the observations were filtered to wind directions which favoured transport from the well pad (contained a westerly component). During both periods, mixing ratios of  $NO_x$  began to increase from 06:00 and remained enhanced throughout the day before declining into the evening and overnight. As a result, Fig. 3 suggests that changes in the PBL height have no obvious effect, implying that  $NO_x$  mixing ratios at KM are strongly driven by local emissions rather than meteorology.

Fig. 3 highlights some clear changes in  $NO_x$  across the two monitoring periods. During the pre-operational period, the range over the day was 3-times greater for both NO and  $NO_x$  compared to the baseline phase. This is primarily due to high daytime mixing ratios of NO, leading to an amplified diurnal cycle throughout the pre-operational period. The largest change was observed for early evening NO, where the mean NO mixing ratio at 17:00 increased by 3975 % from 0.4 ppb in the baseline period to 16.3 ppb during the pre-operational period. This is comparable to the morning rush hour peak in NO (approximately 15 ppb) observed in North Kensington, an urban background site, during the ClearFlo campaign in London (Bohnenstengel et al., 2015). During the pre-operational phase, peak values of  $NO_x$  occurred at 08:00 and 15:00 with an obvious dip at

12:00. This is suggestive of anthropogenic activities, where mixing ratios increase in the morning as the working day begins, decline during a break over lunchtime before increasing again in the afternoon as work resumes.

Also of note is the change in the relative contributions of NO and  $NO_2$  to total  $NO_x$ . During the baseline period  $NO_2$  dominated  $NO_x$  mixing ratios, contributing 83 %. However, during the pre-operational period this trend was reversed and NO became the major component of  $NO_x$ , contributing 59 %. This suggests a change in the most prevalent source of  $NO_x$  at KM, specifically an additional source of primary NO close to the monitoring site such that oxidation to  $NO_2$  was yet to occur. The enhanced structure in the diurnal cycle and change in the predominant component of  $NO_x$  is strong evidence that the pre-operational period had a measurable effect on ambient  $NO_x$  mixing ratios at KM. Additionally, Purvis et al. (2019) reported a 4-fold and 2-fold increase in the annual means of NO and  $NO_x$  from 2016 to 2017, respectively, showing that the pre-operational period led to a significant deterioration in the overall air quality at KM.

The RF models for each pollutant were used to predict the BAU values during the pre-operational period, which ran from 19th September 2017–1st February 2018. Whilst only the baseline data was used for model training, the entire data set was predicted using all available meteorological data as inputs to the RF models. Fig. 4 shows the daily mean time series for observed and predicted mixing ratios between 2016 and 2019 at KM. As expected, the measured and predicted values strongly agree during the baseline phase of monitoring since this data was used to grow and train the RF models. Discrepancies during this period arise when “spikes” occur in pollutant mixing ratios. In part, this is because the models here are regression models and every prediction is an average (mean) of 300 predictions from 300 trees. As a result, the models have a limited ability to capture minima and maxima in pollutant mixing ratios. Significant deviations between the predicted and observed values begin to appear at the beginning of the pre-operational period. Measured  $NO_x$  values are enhanced relative to the predicted values, whilst the opposite is true for  $O_3$ .

In order to evaluate the change in pollutant mixing ratios and to understand the predictions made by the RF models, the general meteorological conditions during the pre-operational period must be considered. Fig. 5 shows the average meteorological variables during the equivalent of the pre-operational period (19th September–1st February) for each year of monitoring at KM. Crucially, the prevailing wind direction was consistently from the west or south west across all years, meaning the monitoring station was ideally located to detect the effect of the activity on site. The

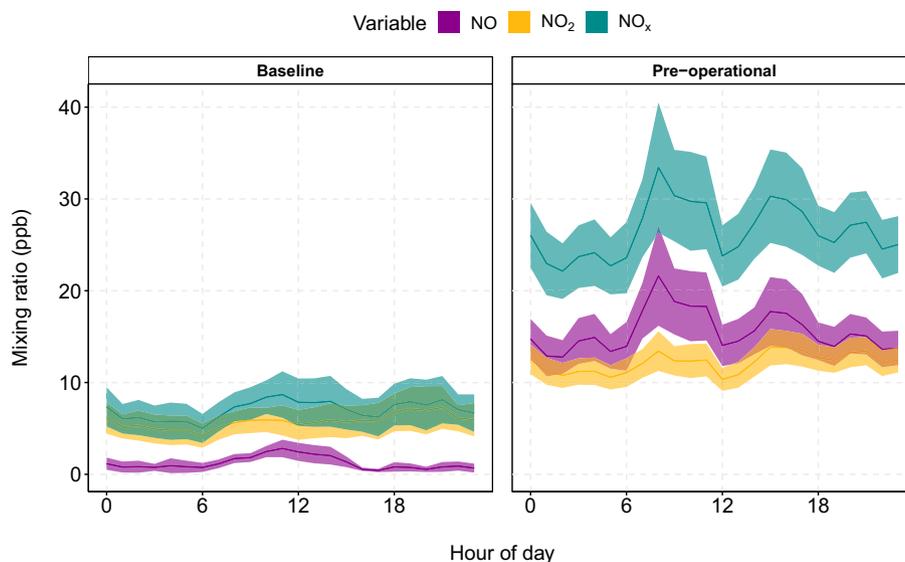


Fig. 3. Average diurnal mixing ratios of NO,  $NO_2$  and  $NO_x$  during westerly winds throughout the baseline and pre-operational periods at KM. Baseline data was filtered between 19th September–1st February for each year. The shaded areas represent the 95 % confidence intervals.

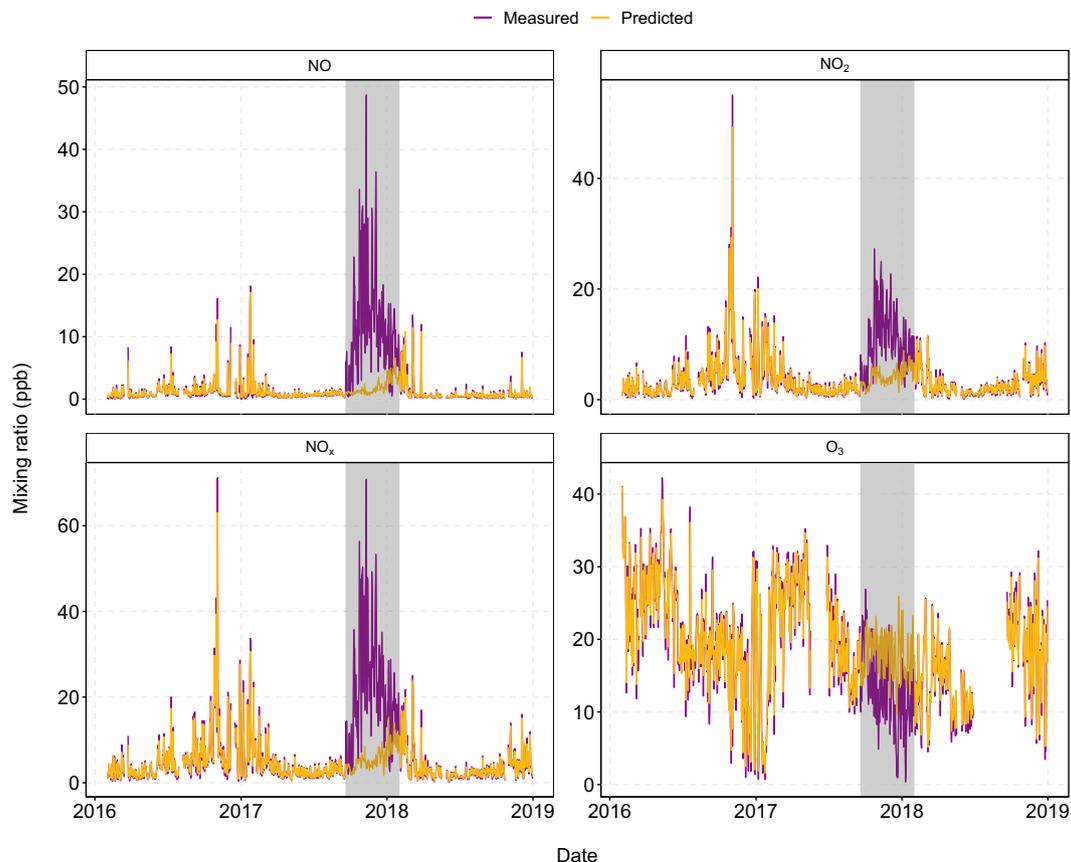


Fig. 4. Measured and predicted daily mean mixing ratios of pollutants at KM. The grey shaded area represents the pre-operational period.

air pressure was lowest during 2017–2018 with a mean value of 1008 mbar compared to 1017 mbar in the previous year. Low air pressure systems generally lead to wet and windy weather conditions. Consequently, this was concurrent with the greatest mean wind speed, which was 40 % higher than the previous year and 60 % higher than the following year. This is expected to lead to lower mixing ratios of pollutants, such as  $\text{NO}_x$ , due to an increase in atmospheric dispersion. Indeed, this is reflected in the model predictions, where the predicted mean  $\text{NO}_x$  during the pre-operational period was 35 % lower than the equivalent period for the previous year. Conversely, the opposite is seen in the measurement data, where total  $\text{NO}_x$  was enhanced 2-fold during the pre-operational period compared to the same

period in the previous year. These results are therefore consistent with the hypothesis that the pre-operational period caused increased mixing ratios of  $\text{NO}_x$ . In terms of  $\text{O}_3$ , the meteorological conditions outlined in Fig. 5 have the opposite effect. Purvis et al. (2019) show that elevated westerly winds generally lead to enhanced  $\text{O}_3$  at KM, therefore predicted  $\text{O}_3$  during the pre-operational period was 19 % higher than the previous year. As was the case for  $\text{NO}_x$ , the observations show the contrary, where  $\text{O}_3$  was 14 % lower than the previous year during the pre-operational period.

To link the divergence from the BAU scenario to the increase in activity due to well preparation, a plot of the increment, defined as observed minus predicted, versus wind direction is shown for  $\text{NO}_x$  and  $\text{O}_3$  in Fig. 6. The plot

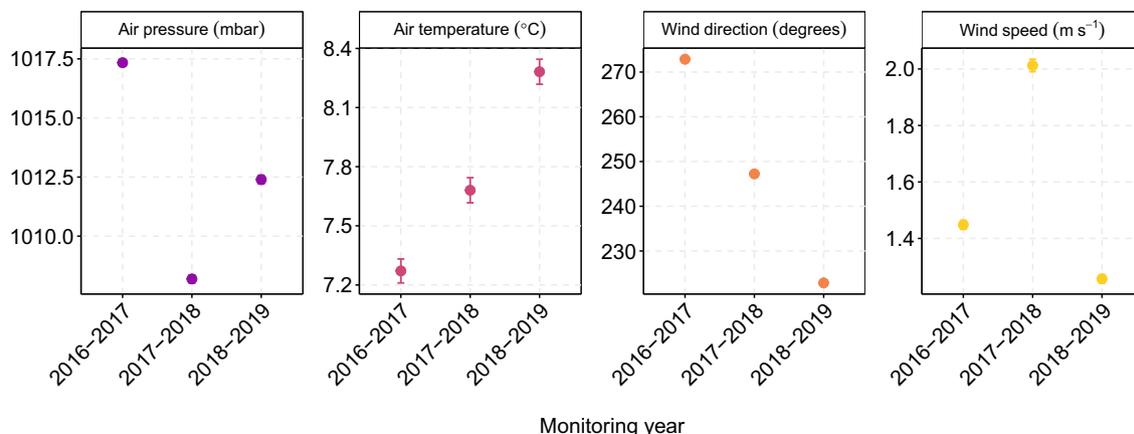
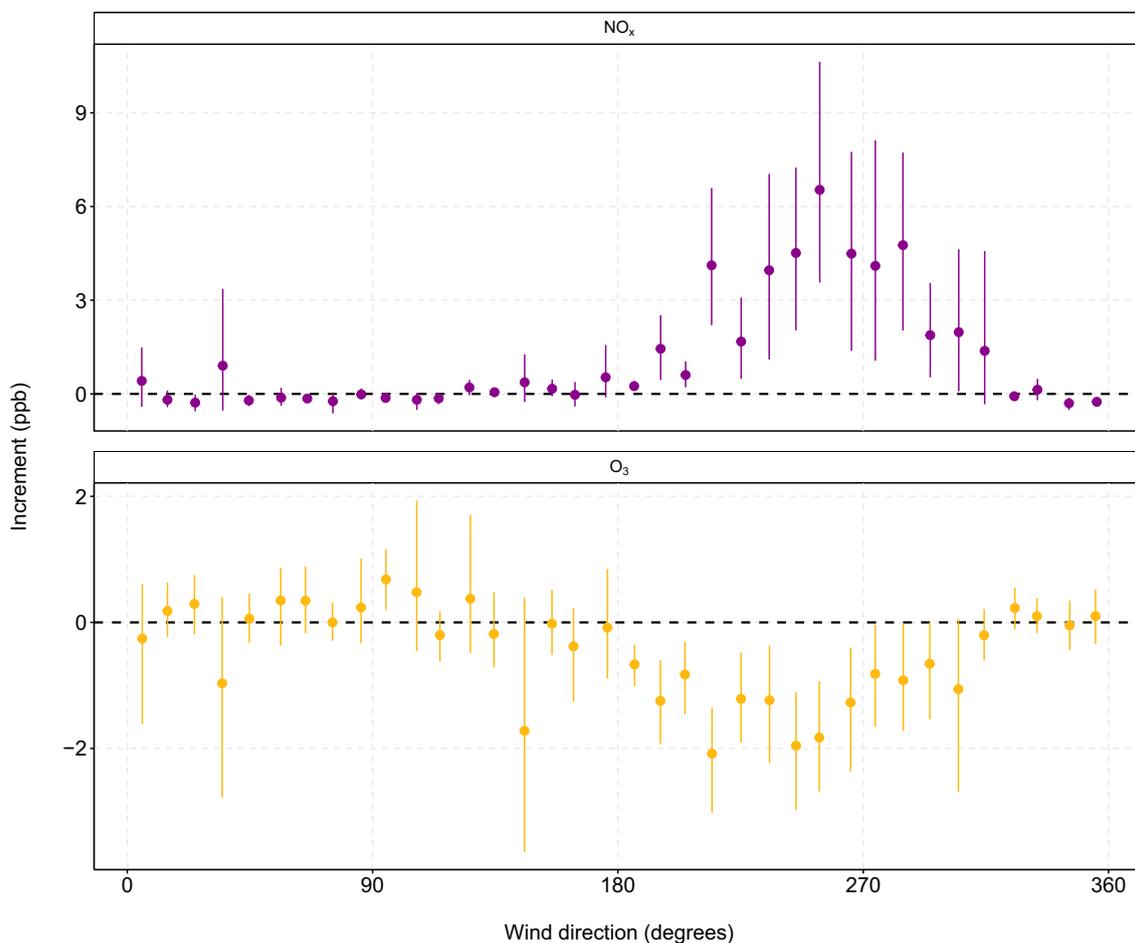


Fig. 5. Means of meteorological variables at KM between 19th September–1st February during each year of monitoring. Error bars show the 95 % confidence intervals in the mean.



**Fig. 6.**  $\text{NO}_x$  (top) and  $\text{O}_3$  (bottom) increment (observed - predicted) by wind direction at KM. Data have been binned into wind direction intervals of 10 degrees and averaged. The error bars represent the upper and lower 95 % confidence intervals in the mean.

shows that the maximum  $\text{NO}_x$  increment is observed for westerly winds ( $\text{NO}_x = 6.48$  ppb), but falls to zero for northerly and easterly winds. Some of the increment shown in Fig. 6 during southerly winds could likely be from idling vehicles associated with the protest campaigning, policing and media presence located outside the site access point on Habton Road. Equivalent plots of NO and  $\text{NO}_2$  (not shown) also displayed an enhancement during westerly winds but the magnitude of the enhancement differed (see discussion below). For  $\text{O}_3$  the trend was reversed such that the largest, negative increments were observed under westerly winds. Fig. 6 is strong evidence that the increment in  $\text{NO}_x$  mixing ratios and concurrent decline in  $\text{O}_3$  is consistent with a change in emission source strength to the west, where the well pad lies.

### 3.3. Quantifying the change due to pre-operational activities

The observed and BAU mixing ratios were used to quantify the air quality impact of the pre-operational period. Table 2 shows the estimated

**Table 2**

Measured and predicted means, deltas (measured-predicted), and percentage change, along with the 95 % confidence intervals for the pre-operational period at KM.

Variable	Measured (ppb)	Predicted (ppb)	Delta (ppb)	% change (%)
NO	$12.6 \pm 0.5$	$1.8 \pm 0.1$	$10.7 \pm 0.5$	$566 \pm 30$
$\text{NO}_2$	$10.6 \pm 0.1$	$4.2 \pm 0.2$	$6.4 \pm 0.3$	$152 \pm 7$
$\text{NO}_x$	$21.6 \pm 0.7$	$5.8 \pm 0.1$	$15.9 \pm 0.7$	$274 \pm 13$
$\text{O}_3$	$12.8 \pm 0.2$	$18.0 \pm 0.1$	$-5.2 \pm 0.3$	$-29 \pm 1.5$

percentage change in the mean mixing ratios of pollutants between 19th September 2017–1st February 2018. The uncertainties in the measured and predicted values represent the 95 % confidence intervals around the mean. Uncertainties in the delta values were found by summing the standard uncertainties for the mean-measured and mean-predicted values in quadrature and subsequently multiplying the result by a coverage factor of  $k = 2$  to give an uncertainty at the 95 % confidence level. The relative uncertainties in the percentage change values were calculated by summing the relative errors of the delta and predicted values in quadrature. The absolute error was then multiplied by a coverage factor of  $k = 2$  to give an uncertainty at the 95 % confidence level. A higher uncertainty, represented as a 95 % confidence interval, was found for the measured values, since there was a much larger variation in mixing ratios compared to the predicted values. For example, measured  $\text{NO}_x$  mixing ratios ranged from 0.5 to 231 ppb, whereas predicted values of  $\text{NO}_x$  only ranged between 1.3 and 20.6 ppb.

Table 2 shows the greatest change was observed for NO, which increased approximately 7-fold, by 566 %, compared to the BAU scenario. The associated increase in  $\text{NO}_2$  was much less (152 %), which is to be expected since primary emissions of  $\text{NO}_x$  are predominantly in the form of NO (Department for Environment, Food and Rural Affairs, 2004). The increase in  $\text{NO}_x$  was accompanied by a decrease in  $\text{O}_3$  of 29 %. Since  $\text{O}_3$  and  $\text{NO}_x$  are closely linked through a chemical cycle within the atmosphere, incremental increases in NO lead to the destruction of  $\text{O}_3$  via titration of the two species. Locations with very high  $\text{NO}_x$  emissions generally do not show as large an increase in  $\text{O}_3$  because the source is in very close proximity and NO mixing ratios remain high relative to oxidant mixing ratios (Department for the Environment, Food and Rural Affairs, 2002). This behaviour is

typical of a roadside monitoring site, which in this case is a good parallel for KM since the monitoring station is located on the well pad itself.

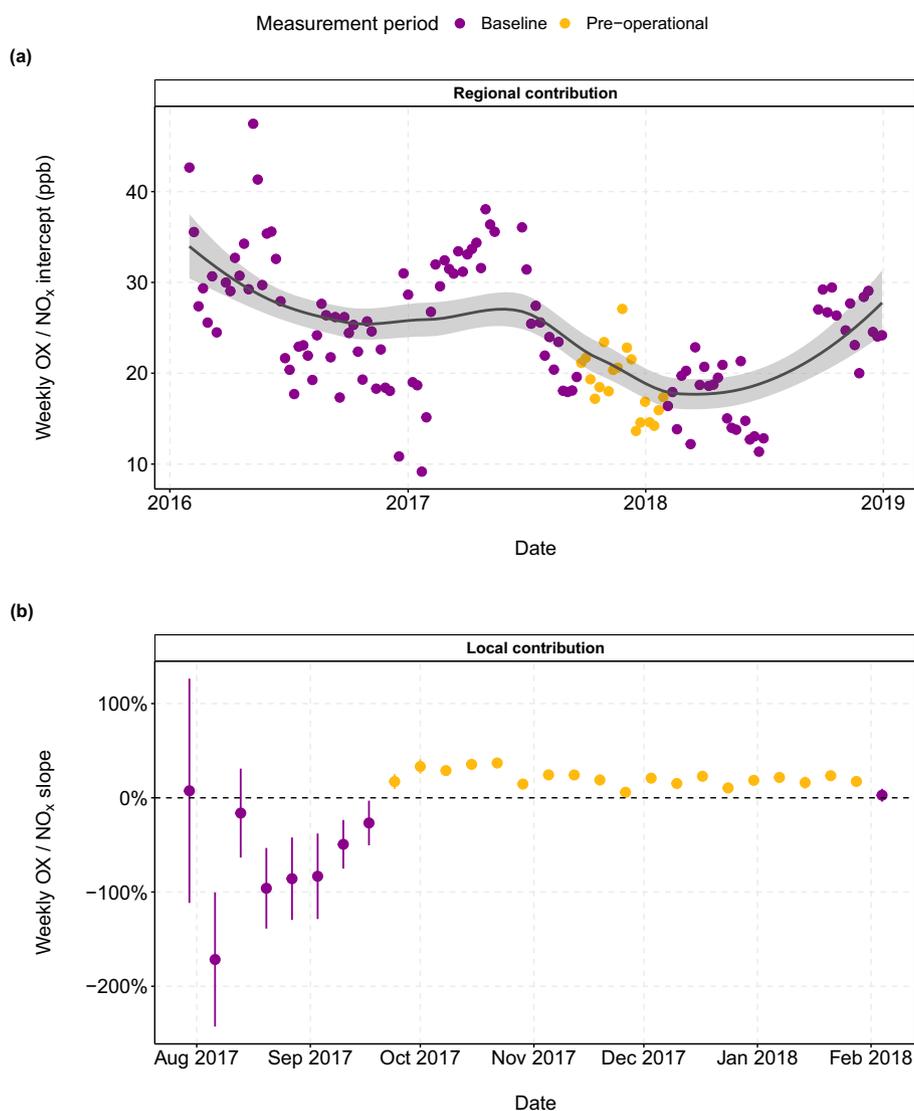
It should also be noted that the KM8 well to be fracked was housed on a pre-existing well pad for conventional gas extraction. However, for brand new wells, the preparation phase would be significantly longer since it would include building the above ground infrastructure, which may require clearing trees, levelling the surface, constructing access roads and laying the well pad itself. Therefore an extended period of preparation is likely to result in larger changes than those reported here. The measured  $\text{NO}_2$  mixing ratios were well below the Air Quality Standard Regulations for the UK, which require that the annual mean concentration of  $\text{NO}_2$  must not exceed  $40 \mu\text{g m}^{-3}$  (19 ppb). However, if pre-operational activities were extended and persisted for an entire year, it's expected that  $\text{NO}_2$  would have exceeded the WHO guidelines for 2021, which set a much stricter recommendation of only  $10 \mu\text{g m}^{-3}$  (5 ppb) (World Health Organization, 2021).

### 3.4. Change in total oxidant

The suppression of  $\text{O}_3$  close to sources of  $\text{NO}_x$  is often accompanied by enhanced levels of  $\text{O}_3$  further downwind. This is due to the oxidation of

$\text{NO}$  to  $\text{NO}_2$  with peroxy radicals and subsequent photolysis of  $\text{NO}_2$  to form  $\text{O}_3$ . Therefore, to account for this photochemistry, the total oxidant ( $\text{OX} = \text{NO}_2 + \text{O}_3$ ) is considered, since production and loss are independent of the chemical coupling that results in the interconversion of  $\text{NO}_2$  and  $\text{O}_3$ . Changes in  $\text{OX}$  reflect the abundance of oxidants and are therefore more representative of the production of oxidant than  $\text{O}_3$  alone (Lu et al., 2010).  $\text{OX}$  can be described in terms of a local,  $\text{NO}_x$ -dependant contribution and a regional,  $\text{NO}_x$ -independent contribution (Clapp and Jenkin, 2001). The regional contribution essentially equates to the regional background level of  $\text{O}_3$ , whereas the local contribution correlates with the level of primary pollution and essentially represents the fraction of directly emitted  $\text{NO}_2$ . The individual contributions to  $\text{OX}$  can be quantified from an  $[\text{OX}]$  vs.  $[\text{NO}_x]$  plot, where the slope obtained from a linear regression represents the local  $\text{OX}$  contribution, whilst the intercept represents the regional contribution (Clapp and Jenkin, 2001).

An additional RF model was constructed for  $\text{OX}$  ( $R^2 = 0.89$ ,  $\text{MSE} = 8.44$  ppb) to assess the change in total oxidant as a result of pre-operational activities at KM. Performing an identical analysis to that for  $\text{NO}$ ,  $\text{NO}_2$ ,  $\text{NO}_x$  and  $\text{O}_3$  yielded a 9 % increase in  $\text{OX}$  relative to the BAU scenario. Fig. 7 shows the local and regional contributions to  $\text{OX}$



**Fig. 7.** (a) Weekly total oxidant ( $\text{OX}$ ;  $\text{NO}_2 + \text{O}_3$ )/ $\text{NO}_x$  intercept at KM calculated using linear regression. The solid line represents a loess smooth fit to the data, and the shaded region represents the 95 % confidence interval around the smooth. (b) Weekly total oxidant ( $\text{OX}$ ;  $\text{NO}_2 + \text{O}_3$ )/ $\text{NO}_x$  slope at KM. Error bars represent the 95 % confidence intervals of the slope estimate.

obtained from the weekly regression plot of [OX] vs. [NO<sub>x</sub>]. Fig. 7a shows that the regional component (intercept) of OX is consistent with the expected seasonal cycle of O<sub>3</sub>, where O<sub>3</sub> generally reaches a minimum during the winter months. Throughout the pre-operational period, the regional contribution to OX followed a declining trend as it approached a minimum and does therefore not account for the observed increase in OX.

Fig. 7b shows the local contribution to OX between August 2017–February 2018. Throughout the baseline period, the fraction of NO<sub>x</sub> directly emitted as NO<sub>2</sub> (f-NO<sub>2</sub>) was negligible, since only secondary NO<sub>2</sub> resulting from the oxidation of primary NO from upwind sources was observed. However, increases in f-NO<sub>2</sub> leading to positive contributions were consistent with the start of the pre-operational period. Throughout the whole of this period, f-NO<sub>2</sub> ranged from 6%–37%, suggesting the increase in OX was driven by changes in primary NO<sub>2</sub> emissions on or near the site. This is likely as a result of the presence of diesel vehicles and generators, which tend to emit a higher f-NO<sub>2</sub> compared to petrol due to diesel emission control technologies such as Diesel Oxidation Catalysts (DOC) (Carslaw et al., 2019). Since access to active O&G sites is exclusively permitted for diesel vehicles, the associated increase in OX is likely to have significant implications on the photochemical production of ozone in regions of hydraulic fracturing.

### 3.5. Site characteristics

Despite the 4-fold increase in total NO<sub>x</sub> during the pre-operational period (Table 2), concentrations were still well below the air quality standards regulations. In order to place the observed changes in concentrations into context, data was compared to that from the Automatic Urban and Rural Network (AURN), the UK's primary air quality monitoring network. The locations of each site are shown in Fig. A.4. Since KM is located in a rural area and not influenced by any single point source, should it be part of the AURN it would likely be classified as a rural background site.

Fig. 8 shows the probability density distribution of NO and NO<sub>2</sub> during both the baseline and pre-operational periods at KM. Baseline observations were filtered to between 19th September–1st February in order to minimise differences due to seasonal factors. Additionally, data from AURN monitoring sites was aggregated to each site classification throughout the same time periods in order to make a robust comparison. Density plots show how the concentrations of pollutants are distributed and give a more detailed indication of where the bulk of measurements lie. This is more useful than using a mean concentration, which can often be skewed by spikes in data. From Fig. 8 it is clear that, as expected, throughout the baseline phase of monitoring, the distributions of NO and NO<sub>2</sub> were most representative of a rural background site. For NO, 93.4% of the measurements were in the interval 0–5 μg m<sup>-3</sup>, almost equivalent to 93.2% for rural background sites. Similarly for NO<sub>2</sub>, 82.8% of data lay in the range 0–10 μg m<sup>-3</sup>, compared to 65% for rural background sites. For comparison, only 21% of data lay within the same range for urban background sites.

For the pre-operational period there is a clear shift in the distribution of both NO and NO<sub>2</sub> concentrations. The NO measurements display a bimodal distribution, in which 31.4% of observations fall into the range 0–5 μg m<sup>-3</sup> and the bulk of observations lie in the range 5–20 μg m<sup>-3</sup>, accounting for 49.8% of observations. This suggests the initial source of NO (likely to be Habton Road) still exists, but that there is also an additional source responsible for higher levels of NO. During this phase, KM is approximately comparable to urban traffic sites, where only 25.1% of data fell into the range 0–5 μg m<sup>-3</sup> and the majority of observations (52.4%) were between 10 and 100 μg m<sup>-3</sup>. There was an evident, albeit smaller change in the distribution of NO<sub>2</sub>. A much broader spread was observed in the data (note the log scale in Fig. 8) with 73.3% of data distributed between 5 and 40 μg m<sup>-3</sup>, compared to 74.3% for urban industrial sites and 55.4% for urban traffic sites.

Fig. 8 indicates that the site characteristics of KM significantly changed following the initiation of unconventional O&G development. Based on NO<sub>x</sub> concentrations, the site transitioned from an air quality

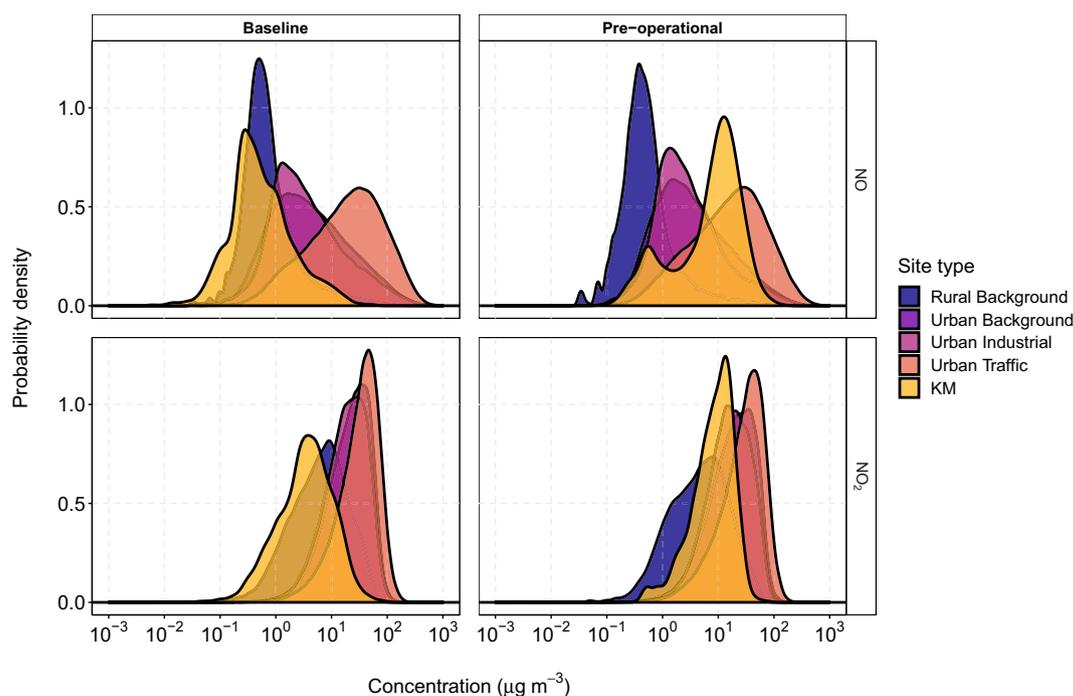


Fig. 8. Normalised probability density functions of hourly mean NO and NO<sub>2</sub> during the baseline and pre-operational periods at KM, compared to UK AURN monitoring sites throughout equivalent periods. Baseline data are filtered to observations between 19th September–1st February for each year.

climatology representative of a rural background site, with relatively clean air typical of the UK regional background, to that more analogous to an urban setting. This could have implications for residents living in the surrounding area of the well site, particularly if the industry were scaled up to facilitate hundreds of wells across the countryside. Should hydraulic fracturing have subsequently taken place, it is expected that emissions would be elevated further above baseline levels. However, it is important to note that the monitoring site was located only 45 m from the well. Emissions of primary pollutants such as NO<sub>x</sub> and VOCs are expected to decrease with increasing downwind distance due to dispersion. Whilst there are a small number of isolated dwellings located within 500 m of the well, local residents are unlikely to live in such close proximity to the well pad and as such will likely be subjected to lower levels of primary pollution than measured here. However, a study in California observed enhanced concentrations of ambient air pollutants within 4 km of pre-production wells and within 2 km of producing wells, suggesting the footprint of emissions from unconventional gas extraction extends far beyond the site itself (Gonzalez et al., 2022).

#### 4. Conclusions

Well pad preparation is a key phase within the shale gas extraction process. Constructing and operating a shale gas well requires a large amount of above ground infrastructure and equipment, which must be transported to the well pad. The resultant traffic load and subsequent on-site activity introduces an additional source of air pollutants to the local environment prior to any hydraulic fracturing. In this work, the impact of the pre-operational phase is investigated through the application of random forest machine learning models to air quality data in the rural village of Kirby Misperton in North Yorkshire.

Extensive baseline monitoring of air pollutants two years prior to the start of shale gas operations enabled the characterisation of the local air quality climatology. The baseline observations were used to predict mixing ratios in the construction of a “business as usual” scenario, which assumed no change in the activity on site. The counterfactual was then compared to the observations, revealing a 274 % increase in NO<sub>x</sub> and concurrent decrease in O<sub>3</sub> of 29 %. Changes in NO<sub>x</sub> were dominated by increases in NO as expected for a traffic-related emission source. However, evaluation of the total oxidant (OX) revealed enhancements of the primary NO<sub>2</sub> fraction (f-NO<sub>2</sub>), which could have negative implications for local public health. Whilst emissions were found to be enhanced significantly above baseline levels, concentrations remained well within UK regulatory limits set for NO<sub>2</sub>. However the concentrations of NO<sub>2</sub> experienced, if sustained year round would have likely been above the 2021 WHO guidelines for NO<sub>2</sub>. Comparison of the data to that from UK AURN monitoring sites dem-

#### Appendix A. Appendix figures

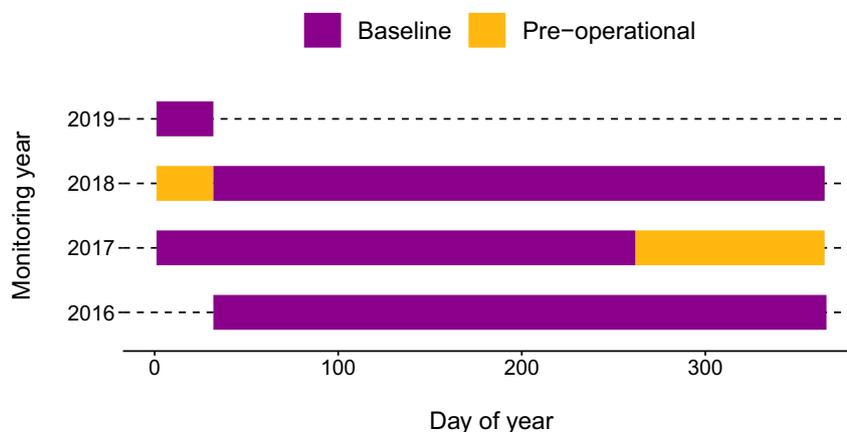


Fig. A.1. Timeline of monitoring periods at KM. The pre-operational phase was active between 19th September 2017–1st February 2018.

onstrated a shift in the chemical environment at KM to that more similar to a suburban city environment in terms of NO<sub>x</sub>. Since no hydraulic fracturing ultimately took place on the site, this work identifies a systematic change in NO<sub>x</sub> due to site preparation in isolation. Often considerations of emissions from unconventional O&G development only emerge once infrastructure is in place and drilling begins. This work therefore exposes a relatively understudied source of emissions from the shale gas industry. Since the desire to reduce dependence on imported energy may refresh interest in domestic gas production in the UK, data that supports the fullest possible assessment of the environmental impacts of activity are vital, and that should include the impacts of pre-operational phases.

#### CRedit authorship contribution statement

**Shona E. Wilde:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Visualization, Writing – original draft, Writing – review & editing. **James R. Hopkins:** Conceptualization, Methodology, Investigation, Supervision, Writing – review & editing. **Alastair C. Lewis:** Conceptualization, Methodology, Investigation, Writing – review & editing. **Rachel E. Dunmore:** Investigation, Project administration, Writing – review & editing. **Grant Allen:** Investigation, Project administration, Writing – review & editing. **Joseph R. Pitt:** Investigation. **Robert S. Ward:** Investigation, Project administration, Writing – review & editing. **Ruth M. Purvis:** Conceptualization, Methodology, Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing.

#### Data availability

Data will be made available on request.

#### 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

The data reported here was collected as part of the Environmental Baseline Project led by the British Geological Survey. This work was funded by the Department for Business, Energy and Industrial Strategy (BEIS; Grant code: GA/18F/017/NEE6617R). The authors would like to thank Third Energy (<https://www.third-energy.com/>) for facilitating the installation of monitoring equipment and the provision of electricity and access during this study. Thanks also go to Sam Wilson for the design and production of the graphical abstract.

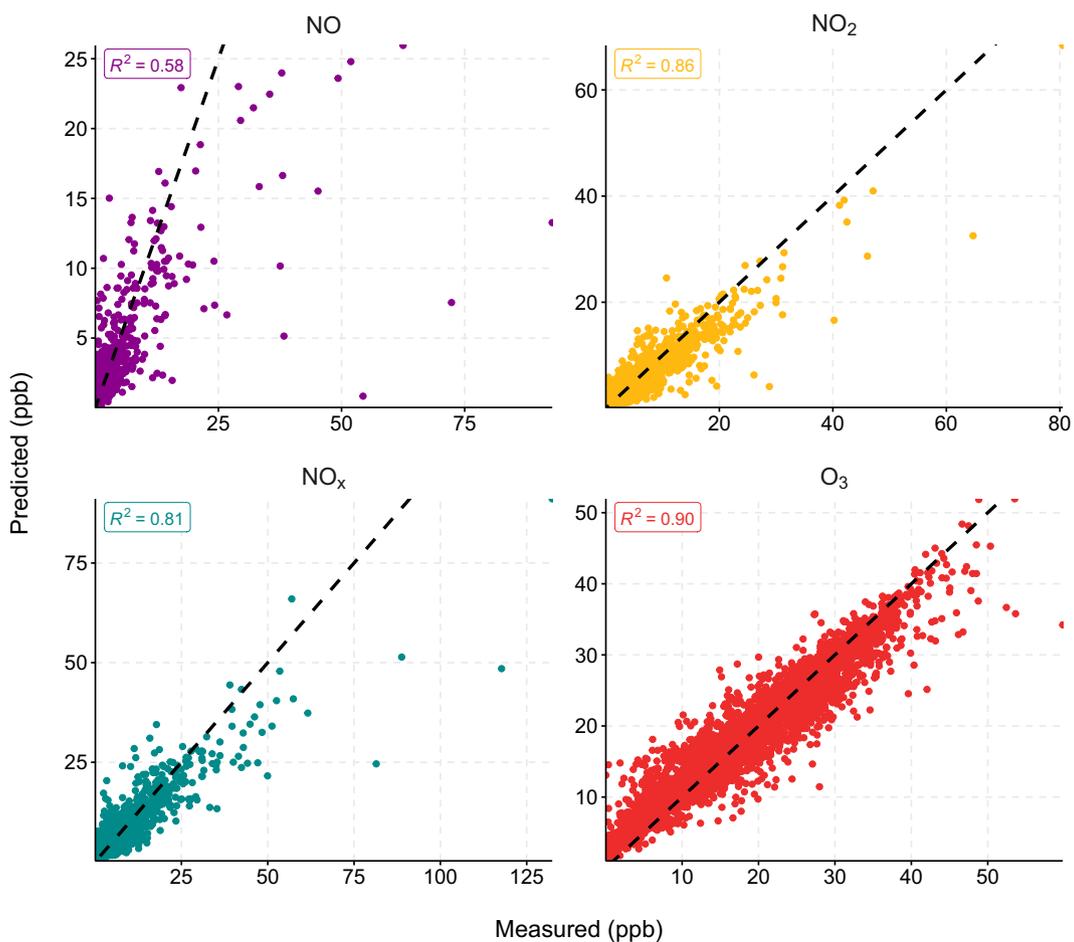


Fig. A.2. Predicted against measured mixing ratios for the testing data set of baseline data for NO, NO<sub>2</sub>, NO<sub>x</sub> and O<sub>3</sub> at KM. The dashed line shows the 1:1 line. The R<sup>2</sup> values are those resulting from a linear fit of the two variables.

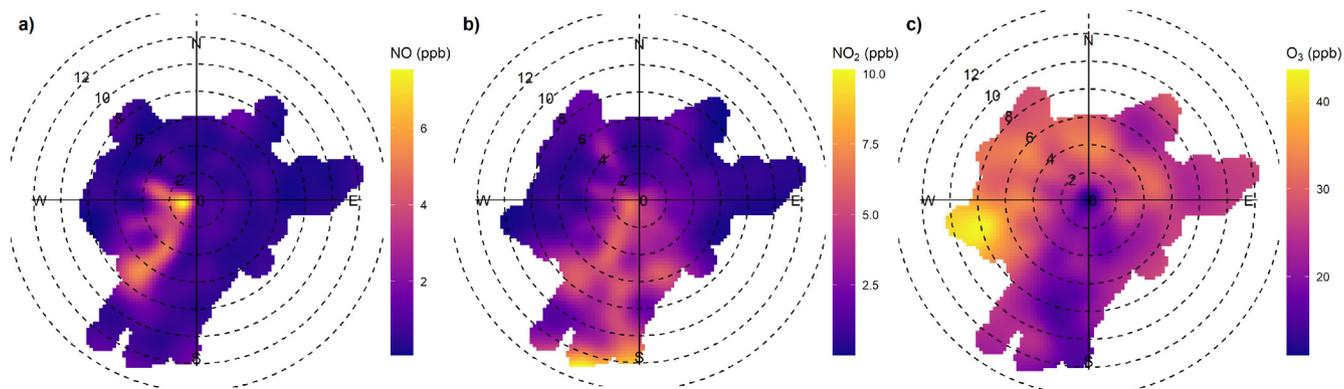


Fig. A.3. Polar plots of NO, NO<sub>2</sub> and O<sub>3</sub> throughout the entire measurement period at KM.

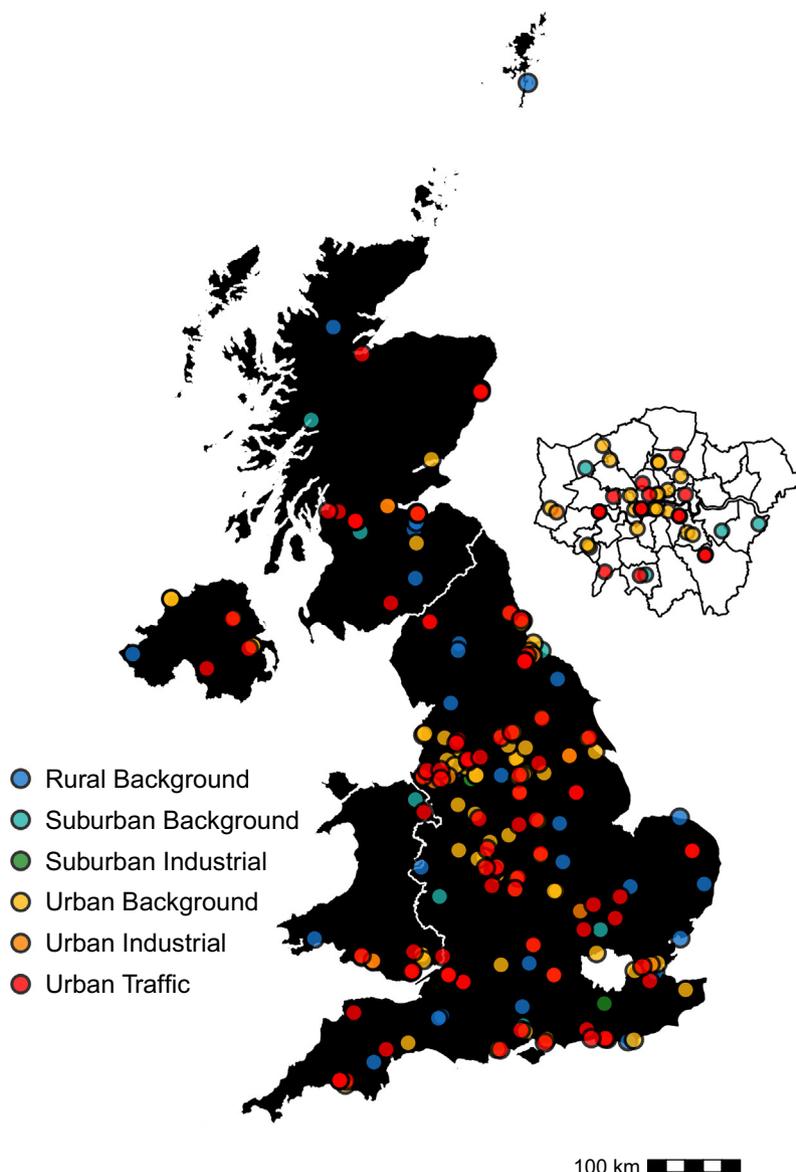


Fig. A.4. The locations of all AURN sites in the UK and their respective classifications. The cut-out shows all London sites.

Appendix B. Appendix tables

Table B.1

RF model performance statistics for NO, NO<sub>2</sub>, NO<sub>x</sub>, O<sub>3</sub> and OX at KM estimated from OOB data (MSE = mean squared error).

Variable	Prediction MSE	R <sup>2</sup>
NO	3.20	0.60
NO <sub>2</sub>	3.82	0.83
NO <sub>x</sub>	10.4	0.76
O <sub>3</sub>	9.57	0.90
OX	8.44	0.89

References

Alvarez, R.A., Pacala, S.W., Winebrake, J.J., Chameides, W.L., Hamburg, S.P., 2012. Greater focus needed on methane leakage from natural gas infrastructure. *Proc. Natl Acad. Sci. USA* 109 (17), 6435. <http://www.pnas.org/content/109/17/6435.abstract>.

Alvarez, R.A., Zavala-Araiza, D., Lyon, D.R., Allen, D.T., Barkley, Z.R., Brandt, A.R., Davis, K.J., Herndon, S.C., Jacob, D.J., Karion, A., Kort, E.A., Lamb, B.K., Lauvaux, T., Maasakkers, J.D., Marchese, A.J., Omara, M., Pacala, S.W., Peischl, J., Robinson, A.L., Shepson, P.B., Sweeney, C., Townsend-Small, A., Wofsy, S.C., Hamburg, S.P., 2018. Assessment of methane emissions from the U.S. oil and gas supply chain. *Science* 361 (6398), 186–188. URL <https://science.sciencemag.org/content/361/6398/186>.

Archibald, A.T., Ordóñez, C., Brent, E., Williams, M.L., 2018. Potential impacts of emissions associated with unconventional hydrocarbon extraction on UK air quality and human health. *Air Qual. Atmos. Health* 11 (6), 627–637 URL <https://doi.org/10.1007/s11869-018-0570-8>.

BBC News, 2019. Kirby Misperton fracking: third energy sells business. URL <https://www.bbc.co.uk/news/uk-england-york-north-yorkshire-48084036>. (Accessed 15 September 2020).

BBC News, 2022a. Fracking ban lifted, government announces. URL <https://www.bbc.co.uk/news/science-environment-62982332>. (Accessed 10 April 2022).

BBC News, 2022b. Fracking will only resume with local consent, PM insists. URL <https://www.bbc.co.uk/news/uk-england-lancashire-63073928>. (Accessed 10 April 2022).

BBC News, 2022c. Ukraine conflict: Petrol at fresh record as oil and gas prices soar. URL <https://www.bbc.co.uk/news/business-60642786>.

Bohnenstengel, S.I., Belcher, S.E., Aiken, A., Allan, J.D., Allen, G., Bacak, A., Bannan, T.J., Barlow, J.F., Beddows, D.C.S., Bloss, W.J., Booth, A.M., Chemel, C., Coceal, O., Di Marco, C.F., Dubey, M.K., Faloon, K.H., Fleming, Z.L., Furger, M., Gietl, J.K., Graves, R.R., Green, D.C., Grimmond, C.S.B., Halios, C.H., Hamilton, J.F., Harrison, R.M., Heal, M.R., Heard, D.E., Helfter, C., Herndon, S.C., Holmes, R.E., Hopkins, J.R., Jones, A.M., Kelly, F.J., Kotthaus, S., Langford, B., Lee, J.D., Leigh, R.J., Lewis, A.C., Lidster, R.T.,

- Lopez-Hilfiker, F.D., McQuaid, J.B., Mohr, C., Monks, P.S., Nemitz, E., Ng, N.L., Percival, C.J., Prévôt, A.S.H., Ricketts, H.M.A., Sokhi, R., Stone, D., Thornton, J.A., Tremper, A.H., Valach, A.C., Visser, S., Whalley, L.K., Williams, L.R., Xu, L., Young, D.E., Zotter, P., 2015. Meteorology, air quality, and health in London: the clearflo project. *Bull. Am. Meteorol. Soc.* 96 (5), 779–804. <https://doi.org/10.1175/BAMS-D-12-00245.1>.
- Boucher, O., Friedlingstein, P., Collins, B., Shine, K.P., 2009. The indirect global warming potential and global temperature change potential due to methane oxidation. *Environ. Res. Lett.* 4 (4), 044007. <https://doi.org/10.1088/1748-9326/4/4/044007>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32. <https://doi.org/10.1023/A:1010933404324> Oct.
- Carslaw, D.C., 2020. Analysis of covid-19 lockdown on UK local air pollution. URL <https://ee.ricardo.com/news/analysis-of-covid-19-lockdown-on-uk-local-air-pollution>.
- Carslaw, D.C., Williams, M.L., Barratt, B., 2012. A short-term intervention study – impact of airport closure due to the eruption of eyjafjallajökull on near-field air quality. *Atmos. Environ.* 54, 328–336. URL <http://www.sciencedirect.com/science/article/pii/S1352231012001355>.
- Carslaw, D.C., Farren, N.J., Vaughan, A.R., Drysdale, W.S., Young, S., Lee, J.D., 2019. The diminishing importance of nitrogen dioxide emissions from road vehicle exhaust. *Atmos. Environ. X* 1, 100002. URL <http://www.sciencedirect.com/science/article/pii/S2590162118300029>.
- Clapp, L.J., Jenkin, M.E., 2001. Analysis of the relationship between ambient levels of o, no and no as a function of no in the UK. *Atmos. Environ.* 35 (36), 6391–6405. URL <http://www.sciencedirect.com/science/article/pii/S1352231001003788>.
- Cole, M.A., Elliott, R.J.R., Liu, B., 2020. The impact of the wuhan covid-19 lockdown on air pollution and health: a machine learning and augmented synthetic control approach. *Environ. Resour. Econ.* (32836865), 1–28. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7416596/>.
- Department for Environment, Food and Rural Affairs, 2004. *Nitrogen Dioxide in the United Kingdom*. May.
- Department for the Environment, Food and Rural Affairs, 2002. *Ozone in the United Kingdom*.
- Derwent, R.G., Middleton, D.R., Field, R.A., Goldstone, M.E., Lester, J.N., Perry, R., 1995. Analysis and interpretation of air quality data from an urban roadside location in Central London over the period from July 1991 to July 1992. *Atmos. Environ.* 29 (8), 923–946.
- Dix, B., de Bruin, J., Roosenbrand, E., Vlemmix, T., Francoeur, C., Gorchoy-Negron, A., McDonald, B., Zhizhin, M., Elvidge, C., Veeffkind, P., Levelt, P., de Gouw, J., 2020. Nitrogen oxide emissions from U.S. oil and gas production: recent trends and source attribution. *Geophys. Res. Lett.* 47 (1), e2019GL085866. <https://doi.org/10.1029/2019GL085866>.
- Edwards, P.M., Brown, S.S., Roberts, J.M., Ahmadov, R., Banta, R.M., Dubé, W.P., Field, R.A., Flynn, J.H., Gilman, J.B., Graus, M., Helmig, D., Koss, A., Langford, A.O., Lefer, B.L., Lerner, B.M., Li, R., Li, S.-M., McKeen, S.A., Murphy, S.M., Parrish, D.D., Senff, C.J., Soltis, J., Stutz, J., Sweeney, C., Thompson, C.R., Trainer, M.K., Tsai, C., Veres, P.R., Washenfelder, R.A., Warneke, C., Wild, R.J., Young, C.J., Yuan, B., Zamora, R., deGouw, J.A., 2014. High winter ozone pollution from carbonyl photolysis in an oil and gas basin. *Nature* 514 (7522), 351–354. <https://doi.org/10.1038/nature13767>.
- Ellsworth, W.L., 2013. Injection-induced earthquakes. *Science* 341 (6142), 1225942. <http://science.sciencemag.org/content/341/6142/1225942.abstract>.
- Field, R.A., Soltis, J., Murphy, S., 2014. Air quality concerns of unconventional oil and natural gas production. *Environ Sci Process Impacts* 16 (5), 954–969. <https://doi.org/10.1039/C4EM00081A>.
- Forster, P.M., Forster, H.I., Evans, M.J., Gidden, M.J., Jones, C.D., Keller, C.A., Lamboll, R.D., Quéré, C.L., Rogelj, J., Rosen, D., Schleussner, C.-F., Richardson, T.B., Smith, C.J., Turnock, S.T., 2020. Current and future global climate impacts resulting from covid-19. *Nat. Clim. Chang.* <https://doi.org/10.1038/s41558-020-0883-0>.
- Gonzalez, D.J.X., Francis, C.K., Shaw, G.M., Cullen, M.R., Baiocchi, M., Burke, M., 2022. Upstream oil and gas production and ambient air pollution in California. *Sci. Total Environ.* 806, 150298. URL <https://www.sciencedirect.com/science/article/pii/S0048969721053754>.
- Goodman, P.S., Galatioto, F., Thorpe, N., Namdeo, A.K., Davies, R.J., Bird, R.N., 2016. Investigating the traffic-related environmental impacts of hydraulic-fracturing (fracking) operations. *Environ. Int.* 89–90, 248–260. URL <http://www.sciencedirect.com/science/article/pii/S0160412016300277>.
- Grange, S.K., Carslaw, D.C., 2019. Using meteorological normalisation to detect interventions in air quality time series. *Sci. Total Environ.* 653, 578–588. URL <http://www.sciencedirect.com/science/article/pii/S004896971834244X>.
- Grange, S.K., Carslaw, D.C., Lewis, A.C., Boleti, E., Hueglin, C., 2018. Random forest meteorological normalisation models for swiss pm trend analysis. *Atmos. Chem. Phys.* 18 (9), 6223–6239. URL <https://acp.copernicus.org/articles/18/6223/2018/>.
- Grange, S.K., Lee, J.D., Drysdale, W.S., Lewis, A.C., Hueglin, C., Emmenegger, L., Carslaw, D.C., 2021. Covid-19 lockdowns highlight a risk of increasing ozone pollution in European urban areas. *Atmos. Chem. Phys.* 21 (5), 4169–4185. URL <https://acp.copernicus.org/articles/21/4169/2021/>.
- Hecobian, A., Clements, A.L., Shonkwiler, K.B., Zhou, Y., MacDonald, L.P., Hilliard, N., Wells, B.L., Bibeau, B., Ham, J.M., Pierce, J.R., Collett, J.L., 2019. Air toxics and other volatile organic compound emissions from unconventional oil and gas development. *Environ. Sci. Technol. Lett.* 6 (12), 720–726. <https://doi.org/10.1021/acs.estlett.9b00591>.
- Helmig, D., Thompson, C.R., Evans, J., Boylan, P., Hueber, J., Park, J.-H., 2014. Highly elevated atmospheric levels of volatile organic compounds in the Uintah basin, Utah. *Environ. Sci. Technol.* 48 (9), 4707–4715. <https://doi.org/10.1021/es405046r>.
- Jaroslowski, J., Pawlak, I., Guzikowski, J., Pietruczuk, A., 2022. Impact of shale gas exploration and exploitation activities on the quality of ambient air - the case study of Wysin, Poland. *Atmosphere* 13 (8).
- King, G.E., 2012. Hydraulic Fracturing 101: What Every Representative, Environmentalist, Regulator, Reporter, Investor, University Researcher, Neighbor and Engineer Should Know About Estimating Frac Risk and Improving Frac Performance in Unconventional Gas and Oil Wells. <https://doi.org/10.2118/152596-MS>.
- Lippmann, M., 1989. Health effects of ozone a critical review. *Japca* 39 (5), 672–695.
- Lu, K., Zhang, Y., Su, H., Brauers, T., Chou, C.C., Hofzumahaus, A., Liu, S.C., Kita, K., Kondo, Y., Shao, M., Wahner, A., Wang, J., Wang, X., Zhu, T., 2010. Oxidant (o + no) production processes and formation regimes in Beijing. *May J. Geophys. Res.* 115 (D7). <https://doi.org/10.1029/2009JD012714>.
- Mediaview, 2012. Mediaview. *Jan Geol. Today* 28 (1), 2–12. <https://doi.org/10.1111/j.1365-2451.2012.00819.x>.
- Priestley, S., 2020. Briefing Paper: Shale Gas and Fracking. House of Commons Library.
- Purvis, R.M., Lewis, A.C., Hopkins, J.R., Wilde, S.E., Dunmore, R.E., Allen, G., Pitt, J., Ward, R.S., 2019. Effects of 'pre-fracking' operations on ambient air quality at a shale gas exploration site in rural North Yorkshire, England. *Sci. Total Environ.* 673, 445–454. URL <http://www.sciencedirect.com/science/article/pii/S0048969719315979>.
- R Core Team, 2020. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Shah, A.S.V., Lee, K.K., McAllister, D.A., Hunter, A., Nair, H., Whiteley, W., Langrish, J.P., Newby, D.E., Mills, N.L., 2015. Short term exposure to air pollution and stroke: systematic review and meta-analysis. *BMJ* 350, h1295. <http://www.bmj.com/content/350/bmj.h1295.abstract>.
- Shaw, J.T., Allen, G., Pitt, J., Mead, M.I., Purvis, R.M., Dunmore, R., Wilde, S., Shah, A., Barker, P., Bateson, P., Bacak, A., Lewis, A.C., Lowry, D., Fisher, R., Lanoisellé, M., Ward, R.S., 2019. A baseline of atmospheric greenhouse gases for prospective UK shale gas sites. *Sci. Total Environ.* 684, 1–13. URL <http://www.sciencedirect.com/science/article/pii/S0048969719323071>.
- Staddon, C., Brown, J., Hayes, E., 2016. Potential environmental impacts of 'fracking' in the UK. *Geography* 101 (2), 60–69. <https://www.jstor.org/stable/26546718>.
- Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. *BMC Bioinf.* 9 (1), 307. <https://doi.org/10.1186/1471-2105-9-307>.
- The Guardian, 2022. Ministers launch fracking study, paving way to end moratorium in England. URL <https://www.theguardian.com/environment/2022/apr/05/ministers-launch-uk-fracking-study-paving-way-to-end-moratorium>.
- U.S. Energy Information Administration, 2020. Annual energy outlook 2020. URL <https://www.eia.gov/outlooks/aeo/>.
- Vengosh, A., Jackson, R.B., Warner, N., Darrah, T.H., Kondash, A., 2014. A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States. *Environ. Sci. Technol.* 48 (15), 8334. <https://doi.org/10.1021/es405118y>.
- Vinciguerra, T., Yao, S., Dadzie, J., Chittams, A., Deskins, T., Ehrman, S., Dickerson, R.R., 2015. Regional air quality impacts of hydraulic fracturing and shale natural gas activity: evidence from ambient VOC observations. *Atmos. Environ.* 110, 144–150. URL <http://www.sciencedirect.com/science/article/pii/S1352231015002800>.
- World Health Organization, 2021. WHO Global Air Quality Guidelines: Particulate Matter (PM<sub>2.5</sub> and PM<sub>10</sub>), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. World Health Organization.
- Zhang, J.J., Wei, Y., Fang, Z., 2019. Ozone pollution: a major health hazard worldwide. *Front. Immunol.* 10. URL <https://www.frontiersin.org/articles/10.3389/fimmu.2019.02518>.
- Zhang, Y., Gautam, R., Pandey, S., Omara, M., Maasackers, J.D., Sadavarte, P., Lyon, D., Nesser, H., Sulprizio, M.P., Varon, D.J., Zhang, R., Houweling, S., Zavala-Araiza, D., Alvarez, R.A., Lorente, A., Hamburg, S.P., Aben, I., Jacob, D.J., 2020. Quantifying methane emissions from the largest oil-producing basin in the United States from space. *Sci. Adv.* 6 (17), eaaz5120. <https://doi.org/10.1126/sciadv.aaz5120>.
- Zheng, H., Kong, S., Chen, N., Yan, Y., Liu, D., Zhu, B., Xu, K., Cao, W., Ding, Q., Lan, B., Zhang, Z., Zheng, M., Fan, Z., Cheng, Y., Zheng, S., Yao, L., Bai, Y., Zhao, T., Qi, S., 2020. Significant changes in the chemical compositions and sources of pm in Wuhan since the city lockdown as covid-19. *Sci. Total Environ.* 739, 140000. URL <http://www.sciencedirect.com/science/article/pii/S0048969720335208>.