Observational data:

There are several networks which measure methane mole fractions and provide observational data. These networks include the UK Deriving Emissions linked to Climate Change (DECC network), the Total Carbon Column Observing Network (TCCON), the COllaborative Carbon Column Observing Network (COCCON)and the Integrated Carbon Observation System (ICOS).1-4 Measurements for these networks are made in a variety of ways **and will be explored at a future date**.

There are several sites in the DECC network for which observational data is available (Table 1). In this summary, data and models for the inlet at a height of 248 m at the DECC Bilsdale site (herein referred to as bsd248) are used to illustrate the work performed so far. The observational methane mole fractions for bsd248 are shown in Figure 1. These observations show that atmospheric methane mole fractions are increasing with time. They also highlight that the mole fraction has a seasonal variation.

Table 1. Sites and inlets in the DECC network for which observational data is available.

|  |  |
| --- | --- |
| **Site** | **Inlets / m** |
| Bilsdale (bsd) | 248, 108 ,42 |
| Heathfield (hfd) | 100, 50 |
| Jodrell bank (jbo) | 45 |
| Ridge hill (rgl) | 90, 45, 0 |
| Lerwick, Shetland islands (sis)  Tall tower Angus (tta) | 2  222, 0 |

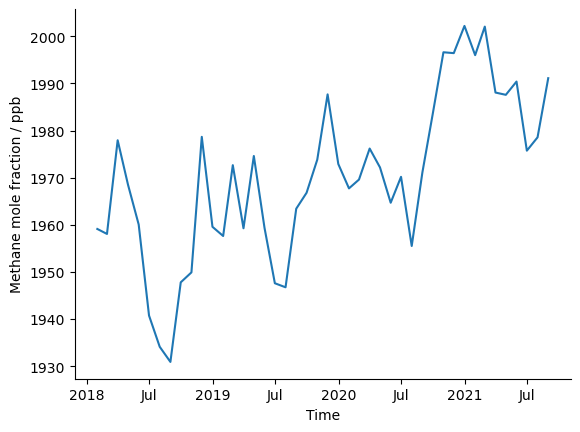


Figure 1. real observed methane mole fractions for bsd248.

Forward models:

To start the project, relatively simple forward models were used to predict methane (CH4) mole fraction observations for different DECC sites (*i.e.* modelled observations were calculated). These modelled observations were then compared to the real observed mole fractions for each site.

To calculate modelled observations, footprint, flux and boundary condition data are required. A footprint is a map which describes the sensitivity of a site to emissions from the surrounding region. Here, footprints calculated using the Met Office's Numerical Atmospheric-dispersion Modelling Environment (NAME) for each site were used **R**. Flux data is an estimate of emissions. Here, data from the EDGAR database was used **R**. Boundary condition data describes the mole fractions at the edges of the modelled domain. Here, the European Centre for Medium-Range Weather Forecasts (ECMWF) Copernicus Atmosphere Monitoring Service (CAMS) camsv19\_daily and camsv22r2\_daily boundary conditions for Europe were used **R**. The domain for these boundary conditions is shown in Figure 2. This data was then combined to give the modelled methane mole fraction observation for a given DECC site.

A map of the world

Description automatically generated

Figure 2. Domain for camsv19\_daily and camsv22r2\_daily boundary conditions.

To calculate modelled mole fraction observations the flux and footprint data are combined to give the modelled pollution events for a given site. This is done by multiplying the emissions and footprint data at a given time and within a given grid cell and then summing over all grid cells as shown in Equation 1.5 The calculated mole fraction data for each time can then be plotted to give the modelled pollution events. The term “pollution events” describes emissions above the baseline atmospheric mole fraction. Therefore, calculating a baseline and adding it to the pollution events gives the modelled observational data for a given site. The baseline is calculated by multiplying the sensitivity of the site to a given a boundary (e.g. sensitivity to the northern boundary) by the mole fraction at that boundary as shown in Equation 2.6 The values obtained for the north, east, south and west boundaries are then combined and multiplied by any losses to give the baseline as shown in Equation 3.6 For these simple models, CH4 was assumed to be an inert species and no loss factor was included when calculating the baseline. For each model, the time period looked at was constricted to the period of overlap between the three data sets and all data was resampled to a monthly average. The time periods looked at in each model are summarised in Table 2. Several sites were not modelled due either to there being no overlap between the data sets (e.g. bsd108 for camsv22r2\_daily) or due to insufficient or missing observational data points (e.g. tta222 and bsd108 respectively for camsv19\_daily ).

(1)

*Where χt is the mole fraction at time t, j is a grid cell within the domain, q are the flux data and ft is the footprint at time t. Equation recreated from intro\_obs\_and\_footprints.ipynb.5*

(2)

*Where ab is the lossless baseline at a given boundary (n/e/s/w), χb is the mole fraction at that boundary and fpb is the sensitivity of the site to that boundary.6*

(3)

*Where ab is the lossless baseline at a given boundary (n/e/s/w), t is time and τ is the lifetime of the species.6*

Table 2. Time period looked at for each model.

|  |  |  |  |
| --- | --- | --- | --- |
| **Boundary conditions** | **Site** | **Start date** | **End date** |
| camsv22r2\_daily | bsd248 | 2018-01-01 | 2021-08-31 |
| camsv22r2\_daily | rgl90 | 2018-01-01 | 2022-12-31 |
| camsv22r2\_daily | hfd100 | 2018-01-01 | 2022-12-31 |
| camsv19\_daily | rgl90 | 2015-01-01 | 2019-12-31 |
| camsv19\_daily | hfd100 | 2015-01-01 | 2019-12-31 |

The modelled pollution events, baseline and mole fraction observations for the inlet at a height of 248 m at the Bilsdale site (herein referred to as bsd248) using the camsv22r2\_daily boundary conditions are shown in Figures 3, 4 and 5 respectively.

A graph of a graph

Description automatically generated with medium confidence

Figure 3. Modelled pollution events for bsd248.

A line graph with numbers and a line

Description automatically generated

Figure 4. Modelled baseline for bsd248.

A line graph with numbers and a line

Description automatically generated

Figure 5. Modelled observed methane mole fractions for bsd248.

For bsd248, the modelled observed mole fractions largely showed the same trends as those seen in the real observed mole fractions (Figure 6). For example, both showed an increase in CH4 mole fraction over time and showed the same pollution peaks (e.g. the three peaks between December 2018 and July 2019). However, despite the trends being similar, the mole fraction values predicted by the model often tended to be over or underpredicted when compared to the real observed values.

A graph with blue and orange lines

Description automatically generated

Figure 6. Modelled and real observed methane mole fractions for bsd248.

To summarise these differences in a more objective manner, the root mean squared error (RMSE), standard deviation (SD) and Pearson correlation coefficient (correlation) between the modelled and real observed mole fractions were calculated. For bsd248 the RMSE was 8.6 ppb, the SD was 8.5 ppb and the correlation was 0.90. Although the correlation was relatively good, the RMSE and SD were high indicating that the model was not of the highest quality. This indicated that the data used could be improved, for example an increase in the quality of the emissions data would likely improve the fit of the model . The stats for models constructed for several other sites using both the camsv19\_daily and camsv22r2\_daily boundary conditions are summarised in Figures 7 and 8 - as for bsd248, most models were not of the highest quality.

A diagram of a graph

Description automatically generated

Figure 7. Taylor diagram summarising stats for models calculated using the camsv19\_daily boundary conditions. Standard deviation is on the x-axis.

A diagram of a mathematical equation

Description automatically generated

Figure 8. Taylor diagram summarising stats for models calculated using the camsv22r2\_daily boundary conditions. Standard deviation is on the x-axis.

Although the simple forward models did not provide highly accurate observed mole fraction estimates, they provided a good introduction to the type of data looked at in this project and served as a starting point for understanding inversion modelling. In inversion modelling, an initial model is constructed, then the emissions estimate is then improved in an iterative fashion until the closest match between the modelled and real observed mole fractions is found.7, 8 Through this, a more accurate top-down emissions estimate can be provided. 7, 8

**References**

1. UK Deriving Emissions linked to Climate Change Network (UK DECC Network), <https://www.bristol.ac.uk/chemistry/research/acrg/current/decc.html>, (accessed 21/10/2024).

2. A. Manning, S. O’Doherty, D. Y. Redington, D. Say, J. Pitt, T. Arnold, C. Rennick, M. Rigby, A. Wisher, A. Wenger and P. Simmonds, *Long-Term Atmospheric Measurement and Interpretation of Radiatively Active Trace Gases*, U. o. Bristol, Bristol, 2022.

3. TCCON Charter, <https://tccon-wiki.caltech.edu/Main/TCCONCharter>, (accessed 21/10/2024).

4. COCCON, <https://www.imk-asf.kit.edu/english/COCCON.php>, (accessed 21/10/2024).

5. Getting started with atmospheric mole fraction data and NAME footprints, <https://github.com/mrghg/undergrad_projects/blob/main/notebooks/intro_obs_and_footprints.ipynb>, (accessed 18/10/2024).

6. OpenGHG modelscenario analyse, <https://docs.openghg.org/api/api_analyse.html>, (accessed 18/10/2024).

7. A. J. Manning, S. O'Doherty, A. R. Jones, P. G. Simmonds and R. G. Derwent, *Journal of Geophysical Research: Atmospheres*, 2011, **116**.

8. M. F. Lunt, A. J. Manning, G. Allen, T. Arnold, S. J. B. Bauguitte, H. Boesch, A. L. Ganesan, A. Grant, C. Helfter, E. Nemitz, S. J. O'Doherty, P. I. Palmer, J. R. Pitt, C. Rennick, D. Say, K. M. Stanley, A. R. Stavert, D. Young and M. Rigby, *Atmos. Chem. Phys.*, 2021, **21**, 16257-16276.