

NAT-MN0Y

2015-16 Natural Sciences Project

Investigating terrestrial biosphere activity and fossil fuel CO₂ emissions with atmospheric potential oxygen measurements at Mace Head, Ireland

Author:
100005300

Supervisor:
Dr Andrew Manning

Acknowledgements

I would like to thank my supervisor Andrew Manning for giving me the resources required to undergo the project and setting aside countless hours to provide the help and guidance. I am also very grateful to Penelope Pickers who continuously answered my questions and actively guided me throughout the project, in particular she helped me get to grips with R and provided complex code which was required, but outside the scope of this project.

For the CO₂ and O₂ data I would like to thank Andrew Manning and the Mace Head station manager Gerry Spain (University of Ireland, Galway, Ireland). I would also like to thank Alison Craggs who ‘cleaned up’ O₂ and CO₂ data up until early February 2015.

For the meteorological data and CO data at Mace Head I thank Michel Ramonet and Simon O’Doherty respectively.

Abstract

Information on the atmospheric molar fraction of fossil fuel derived CO₂ (ffCO₂) can be used to validate emissions inventories, a necessity for climate change mitigation. It can also be used to isolate the CO₂ signal from the terrestrial biosphere (bioCO₂), allowing its activity and function as a source or sink be studied. Current methods to determine the atmospheric ffCO₂ mole use tracers which show some atmospheric relation to fossil fuel burning, unfortunately under many circumstances they are associated with large uncertainties and biases. Atmospheric potential oxygen (APO) has been suggested as a novel ffCO₂ tracer, one which has far less implicating factors. Here APO is found it to be more precise than CO, which is a commonly used tracer; furthermore these results suggest it may also be more accurate.

Deriving the temporally varying ffCO₂ molar fraction at Mace Head, Ireland using APO the bioCO₂ component of the CO₂ signal has been identified. At Mace Head bioCO₂ is found to contribute almost solely to the diurnal fluctuations of CO₂ in September 2014, due to daytime photosynthesis and night-time respiration. Large variations between days in the amplitude of the bioCO₂ diurnal cycle are attributed to air mass history, local meteorological conditions and the planetary boundary layer (PBL) depth. On days where the air mass has a maritime origin the diurnal cycles have a small amplitude, whilst those which have passed over local land masses possess the largest bioCO₂ diurnal cycle. Local meteorological conditions and the PBL depth also appear inherently related to the bioCO₂ signal, and this is supported by strong theoretical framework. However, identifying and quantifying the effects of each of these variables could not be achieved within the scope of this study.

Keywords: CO₂ emissions; APO; radiocarbon, air mass history, photosynthesis, planetary boundary layer

Contents

Contents.....	4
1. Introduction	5
1.1. Climate change	5
1.2. Fossil fuel derived CO ₂	5
1.3. Terrestrial biosphere derived CO ₂	6
1.4. Radiocarbon, a fossil fuel CO ₂ tracer	8
1.5. Alternative ffCO ₂ tracers	9
1.6. Atmospheric Potential Oxygen: a novel ffCO ₂ tracer	10
1.7. Environmental controls on the activity of the biosphere	11
1.8. Vertical mixing	11
1.9. Atmospheric transport	13
1.10. Aims and objectives	13
2. Methodology	14
2.1. The measurement site	14
2.2. Data handling	15
2.3. Baseline generation	17
2.4. Generating the emissions ratios	18
2.5. Calculating the ffCO ₂ and bioCO ₂ mixing ratios	18
2.6. Uncertainty production	19
2.7. Data analysis	20
2.8. Determining the air mass history	21
3. Results and Discussion	22
3.1. Baseline selection	22
3.2. Uncertainty analysis	24
3.3. CO ₂ seasonal variability	27
3.4. CO ₂ diurnal variability	28
3.5. Meteorological effects on the local biosphere	29
3.7. The planetary boundary layer	33
3.8. Effect of meteorological variables on the PBL	35
3.9. The non-correlated meteorological parameters	35
3.10 Air mass history – wind speed and direction	37
3.11. Air mass history – back trajectories	38
3.12. Caveats in the ffCO ₂ methodology and areas for future research	41
3.13. Further areas of research	42
4. Conclusion	43
References	45
5. Appendix	51

1. Introduction

1.1. Climate change

Between the late 19th century and 2000 the Earth's global mean surface temperature (GMST) increased by ~ 0.7 °C, since the mid-20th century at least it is extremely likely (> 95 % probability) that this increase has been caused by anthropogenic emissions of greenhouse gases (IPCC, 2013). It is clear from atmospheric CO₂ measurements which span back to 1958 at Mauna Loa, Hawaii (Figure 1) that its atmospheric molar fraction has increased dramatically, primarily due to fossil fuel combustion, as well as cement production and land clearing (Schimel, *et al.*, 2001). CO₂ could in fact responsible for 90 % of the anthropogenic contribution to climate change (Friedlingstein, *et al.*, 2014).

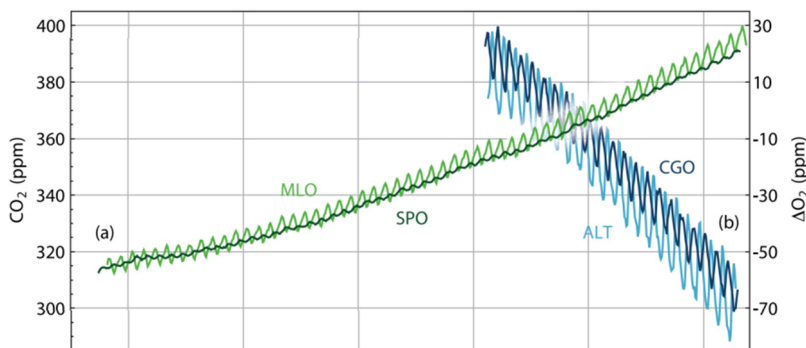


Figure 1 - Anthropogenic CO₂ emissions are driving the atmospheric increase of CO₂. As CO₂ is increasing, a synonymous decrease in O₂ can be seen. A clear seasonal cycle can be seen in both the O₂ and CO₂ data, where CO₂ increases in northern hemisphere winter, and decreases in summer (IPCC, 2013, fig.6.3).

As part of the conference of parties twenty-first session (COP21) Paris climate agreement the GMST is to be limited to a maximum of 2 °C increase, with efforts being pursued to further limit this to 1.5 °C, relative to the pre-industrial GMST (UNFCCC, 2015). Without increasing mitigation efforts it is inevitable that we will exceed the

prescribed GMST limits, maybe as soon as in thirty years' time (Friedlingstein, *et al.*, 2014). It is therefore vital that anthropogenic greenhouse gas emissions are reduced dramatically, in particular CO₂, as it is having the most prominent effect.

1.2. Fossil fuel derived CO₂

Estimates of the atmospheric molar fraction of fossil fuel derived CO₂ (ffCO₂) can be used to quantify fossil fuel emissions, a necessity for climate change policy. Furthermore, ffCO₂ can be used to isolate the terrestrial biosphere derived CO₂ (bioCO₂), allowing its activity to be studied (Turnbull, *et al.*, 2006, Van der Laan, *et al.*, 2010).

It is protocol for countries report their emission statistics biennially, to ensure sufficient mitigation is occurring (UNFCCC, 2011). Currently, statistical inventories are used to report these emissions, using data on energy production, usage and trade from a host of sources such as the United Nations and the International Energy Agency (IEA) as well as archived data (Andres, *et*

al., 2011). This methodology which has been termed the ‘bottom-up’ approach, and reports suggest that their methodology had led to uncertainties in excess of 50 % in some national inventories (Nisbet and Weiss, 2010). Biases may be included in these inventories by incorrect data sources on emissions, whilst the extrapolation of inventories which is often required to obtain the most up to date data also incurs added uncertainty (Levin, *et al.*, 2003, Turnbull, *et al.*, 2006). On top of this, the value of emissions reductions in carbon markets and political pressure generates incentives for the emission data sets to be intentionally skewed (Nisbet and Weiss, 2010).

The top-down approach which uses atmospheric measurements can be valuable for validating these emissions inventories. Once the atmospheric molar fraction of ffCO_2 has been identified (the methodology of which will be discussed later), gridded spatial and temporal surface emission fluxes can be quantified by running atmospheric transport models in an inverse mode, a process which is effective at global to regional scales (Manning, 2011). Unfortunately, there are large sources of uncertainty within these inversion models, and as to be discussed throughout this report the methodology for ffCO_2 quantification (Manning, 2011). However, using both the top-down and bottom-up methods in conjunction with each other for cross-validation would surely increase the accuracy of emissions reporting.

1.3. Terrestrial biosphere derived CO_2

As shown in Figure 2, carbon is exchanged between various reservoirs in the global carbon cycle, importantly for this report, CO_2 is transferred to the atmosphere via fossil fuel burning, ocean transfer and terrestrial biosphere efflux. The total atmospheric CO_2 (obsCO_2) therefore contains a background component (bgCO_2), ffCO_2 component and the terrestrial biosphere component (bioCO_2), the oceanic component is expressed within bgCO_2 (Gamnitzer, *et al.*, 2006). Using this information it is easy to isolate the bioCO_2 component when ffCO_2 and bgCO_2 are known, bgCO_2 can be derived by various methods as to be explained later.

Quantification of bioCO_2 can have many functions, one such use is to assess the role of the terrestrial biosphere as a carbon source/sink (Gamnitzer, *et al.*, 2006, Levin, *et al.*, 2003). Up until now the ocean and the terrestrial biosphere (probably since the 1980s) have acted as net carbon sinks, removing CO_2 from the atmosphere, leaving only an airborne fraction of ~40 % (Knorr, 2009). The response of carbon sources or sinks and subsequently the CO_2 airborne fraction to increasing atmospheric CO_2 , known as the carbon cycle feedback is one of the largest uncertainties faced when predicting the Earth’s future climate (Friedlingstein, *et al.*, 2014). Evidence does however suggest that both the oceanic and terrestrial carbon sink are weakening,

meaning a larger proportion of emitted CO₂ will stay airborne (Keeling, *et al.*, 1995) (Canadell, *et al.*, 2007). An analysis of eleven coupled climate-carbon cycle models found that eight models attributed most of the carbon-feedback sensitivity to the terrestrial biosphere, whilst three attributed most to the oceanic sink (Friedlingstein, *et al.*, 2006).

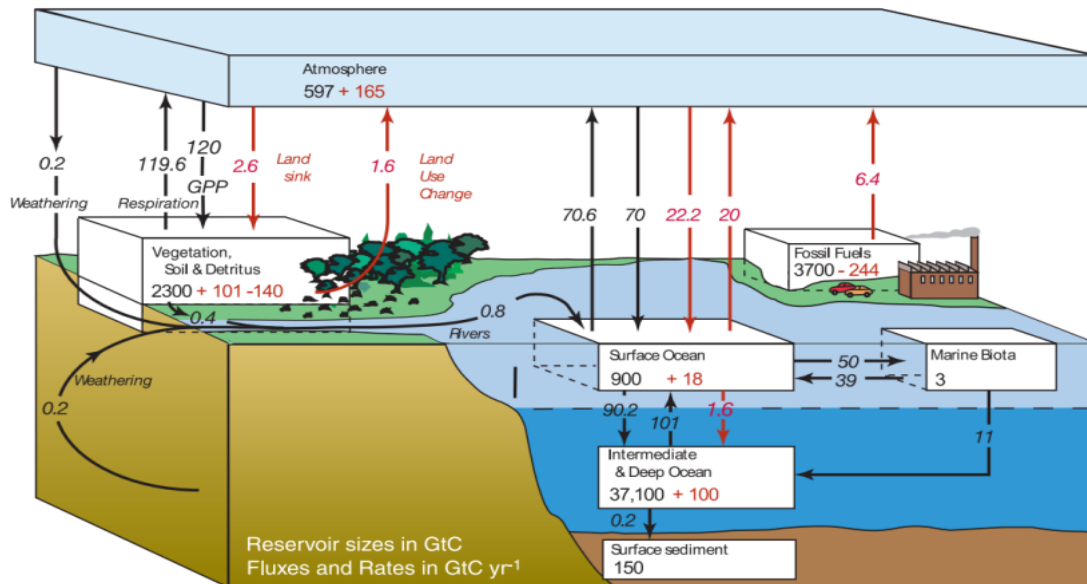


Figure 2 – The global carbon cycle for the 1990s. The pre-industrial fluxes are in black, whilst the anthropogenic fluxes are in red (IPCC, 2007, fig.7.3)

It is clear that understanding the dynamics of the terrestrial biosphere is necessary to further constraining our understanding of carbon cycle feedbacks, and this is invaluable to predict future climate change. The transient climate response to cumulative carbon emissions (TCRE) is the gradient of the response of GMST to an increase in CO₂ emissions; it can be used to set a cumulative emissions budget which must not be exceeded if GMST decrease is to stay below 2 °C (Friedlingstein, *et al.*, 2014). The equilibrium climate sensitivity (ECS) on the other hand is the GMST increase associated with a doubling of CO₂ concentrations in the atmosphere, a vital input for climate models (Friedlingstein, *et al.*, 2006). Both TCRE and ECS are prone to uncertainties from the carbon cycle feedback, and as such can only be constrained between 0.8 – 2.5 °C per 1000 GtC and 1.5 – 4.5 °C respectively (IPCC, 2013).

Diurnal and seasonal cycles of atmospheric CO₂, the latter of which can be seen in Figure 1, are predominantly forced by the terrestrial biosphere. The seasonal cycle occurs as the growing season in Northern hemisphere summer causes a mass CO₂ drawdown from the atmosphere, whilst the opposite is seen during winter (Keeling, *et al.*, 1996). Circadian-like diurnal cycles can also be seen during the summer, driven by daytime photosynthesis and subsequent net CO₂ drawdown during the day followed by biotic respiration at night and CO₂ efflux (Miyaoka, *et al.*, 2007). BioCO₂ data could be vital for understanding the behaviour of these cycles of carbon flux,

and validate the results of process models (Molina-Herrera, *et al.*, 2015). Using bioCO₂ values as opposed to obsCO₂ is advantageous when studying the terrestrial biosphere as it removes the fossil fuel signal, which has its own global emissions seasonal cycle with a peak in January followed by a trough of emissions in August (Rotty, 1987). Although generally the ffCO₂ cycle does not have a noticeable effect on the total atmospheric mixing ratio of CO₂, it has been suggested that it could be noticeable at certain measurement sites (Andres, *et al.*, 2011). Dependent on location there may also be an ffCO₂ diurnal cycle (Hernandez-Paniagua, *et al.*, 2015, Lopez, *et al.*, 2013).

1.4. Radiocarbon, a fossil fuel CO₂ tracer

The first step to quantifying bioCO₂, is to accurately determine ffCO₂. Due to large natural fluxes of CO₂ and atmospheric mixing which integrates emissions in space and time, unique ffCO₂ tracers must be used, and CO₂ increases cannot just be measured (Van der Laan, *et al.*, 2010). Radiocarbon (¹⁴C) is a highly desirable tracer, described by Bozhinova *et al.* (2014) as the ‘gold standard’ for ffCO₂ quantification. Atmospheric ¹⁴C was predominantly produced in the 1950s and 1960s by nuclear bomb testing. Now is only produced in small quantities by nuclear power production and cosmic rays interacting with neutrons and nitrogen, within the atmosphere it is well-mixed (Graven and Gruber, 2011). It can be used to quantify ffCO₂ as it has a half-life of just 5568 years, meaning it is inexistent in fossil fuels (Stuiver and Polach, 1977). Anthropogenic burning of fossil fuels therefore leads to a reduction in the atmospheric ¹⁴C:¹²C ratio by dilution (which is expressed as Δ¹⁴C, a deviation from an oxalic acid in ‘per mil’) (Suess, 1955). Using equation 1, ffCO₂ can be quantified where obsΔ¹⁴C and bgΔ¹⁴C represent the total and background Δ¹⁴C mole fractions (Levin, *et al.*, 1989). This technique has been used extensively i.e. (Levin, *et al.*, 2003, Van der Laan, *et al.*, 2010).

$$ffCO_2 = obsCO_2 \times \frac{bg\Delta^{14}C - obs\Delta^{14}C}{bg\Delta^{14}C + 1000} \quad 1$$

Grab sample measurements of ¹⁴C can be made by using accelerated mass spectrometry to provide point sources of point data (Turnbull, *et al.*, 2006), or alternatively integrated samples which have a resolution of a week or more can be produced using a CO₂ counter (Schoh, *et al.*, 1980). Unfortunately ¹⁴C measurements are expensive to make and labour intensive, making it logistically impractical to produce data at the required spatial and temporal resolution. In the integrated carbon observation system for example, each measurement site produces just 50 ¹⁴C measurements per year (Vardag, *et al.*, 2015). Advances are being made to couple gas chromatography with continuous flow accelerator mass spectrometry to obtain a higher

frequency of ^{14}C measurements; this however is not yet a readily available technology (Vardag, *et al.*, 2015). Perhaps the only issue facing the accuracy of using ^{14}C as an ffCO_2 tracer is the bias which occurs as a result of ^{14}C release from nuclear power plants, acting to dilute the calculated ffCO_2 values. Those power plants which use pressurised water reactors release what can be considered almost a negligible amount of ^{14}C , whilst gas cooled reactors can produce the most ^{14}C per unit of energy production (up to 5.5 TBq GWa^{-1}), along with nuclear fuel reprocessing plants (Graven and Gruber, 2011). The British Isles are prone to large nuclear bias due to several gas-cooled reactor nuclear power plants and the Sellafield reprocessing site. Over the English channel the offset in $\text{ffCO}_2(^{14}\text{C})$ can be as large as 260 %, in Ireland specifically, the nuclear bias is expected to exceed 20 %, (Graven and Gruber, 2011). This makes it almost impossible to use ^{14}C as an accurate ffCO_2 tracer from most locations in the British Isles.

1.5. Alternative ffCO_2 tracers

To produce ffCO_2 data at a high temporal resolution, various alternative tracers have been considered. Three rules must be adhered: (i) have a source that is uniquely related to that of ffCO_2 , (ii) be easy to measure precisely, a high temporal resolution of measurements will provide added benefit and (iii) it should be conserved in the atmosphere or its sink understood well (Gamnitzer, *et al.*, 2006).

Carbon monoxide (CO) is considered to be a candid tracer as fossil fuel combustion where CO_2 is an end product, always forms some amount of CO (Levin and Karstens, 2007). CO measurements can also be made cheaply at a high temporal resolution using commercially available analysers (Bozhinova, *et al.*, 2014). ffCO_2 can be quantified by using CO in accordance with equation 2, where obsCO is the total CO, bgCO the atmospheric background CO and R_{CO} the emissions ratio of $\text{CO}:\text{CO}_2$ associated with the combustion of fossil fuels (Gamnitzer, *et al.*, 2006).

$$\text{ffCO}_2 = \frac{\text{obsCO} - \text{bgCO}}{R_{\text{CO}}} \quad 2$$

Although widely used i.e. (Lopez, *et al.*, 2013, Turnbull, *et al.*, 2006, Vardag, *et al.*, 2015) CO violates rules (i) and (iii) as it has sources which are not related to fossil fuel combustion, nor is it conserved in the atmosphere. CO is produced in large quantities by the oxidation of volatile organic compounds and various non-fossil fuel hydrocarbons (such as methane), whilst surface fluxes from the soil and oceans contribute greatly to the atmospheric CO budget, in fact these are responsible for the production of a larger amount of CO than fossil fuel combustion (Granier, *et al.*, 2000). To make matters worse, the natural CO sources are clouded with large amounts of

uncertainty (Granier, *et al.*, 2000). Most importantly however, biomass burning and the combustion of solid biofuels also produce CO, the latter being CO's largest limitation, releasing a sizeable 30 CO ppb CO₂ ppm⁻¹ (Vardag, *et al.*, 2015). As biofuels may well become an important fuel source required to mitigate climate change the CO method's sensitivity to their use is a real hindrance (Sims, *et al.*, 2006). Oxidation with hydroxyl radicals act as the largest atmospheric CO sink, this process is subject to diurnal and seasonal cycles which still are not completely understood; CO stability is therefore highly variable, ranging from a few weeks to a year (Gamnitzer, *et al.*, 2006, Granier, *et al.*, 2000). R_{CO} can be offset by up to 20 % due to uncertainties in the CO sinks and sources (Gamnitzer, *et al.*, 2006, Turnbull, *et al.*, 2006).

Other tracers such as SF₆ and C₂H₂ have also been considered, but due to significant uncertainties in their emissions, none can supersede CO (Turnbull, *et al.*, 2006). In order to obtain the most accurate ffCO₂ data possible at a high temporal resolution ¹⁴CO₂ measurements (integrated and grab samples) can be used to regularly calibrate the ffCO₂(CO) results, termed to ¹⁴C-calibrated CO method (Vogel, *et al.*, 2010). In this process the ffCO₂(¹⁴C) results are used at regular time intervals to set the 'correct' R_{CO}, obviously this can only be used in regions with little nuclear influence (Levin and Karstens, 2007).

1.6. Atmospheric Potential Oxygen: a novel ffCO₂ tracer

There is clearly a need for alternative ffCO₂ tracers to reduce the uncertainty in ffCO₂ measurements and move away from our reliance on expensive and tedious ¹⁴C measurements. Atmospheric potential oxygen (APO) is an atmospheric tracer, where $APO \approx [O_2] + 1.1 \times 4.8 \times [CO_2]$, APO and oxygen (O₂) are in "per meg" and CO₂ in ppm, the 4.8 converts CO₂ into "per meg" whilst the 1.1 represents the oxidative ratio (O₂:C) of biomass (Stephens, *et al.*, 1998). The oxidative ratio term ensures that APO is conservative with respect to the biosphere, essentially representing the ratio of O₂ consumed to CO₂ emitted for respiration and vice versa for photosynthesis (Keeling, 1988). Changes in the APO values are in fact only caused by the air-sea exchange of CO₂, O₂ and N₂ (the oceanic influence), and unfortunately the combustion of fossil fuels or liquid or gaseous biofuels (Keeling, 1988, Stephens, *et al.*, 1998)(Pickers, personal communication). On short time scales at least the oceanic influence on APO is negligible, Hamme & Keeling (2008) observed an interannual oceanic influence of just 5 – 10 per meg. APO can therefore be considered to fulfil rule (i) far better than CO as the source of change (for short time scales at least) can be attributed to far fewer factors.

Combustion of fossil fuels and liquid or gaseous biofuels cause negative excursions in the APO data compared to the background because these fuels sources have a higher oxidative ratio than

the biosphere, ranging from 1.17 for coal to 1.95 for gaseous fuels (Steinbach, *et al.*, 2011). The use of APO as a tracer for ffCO₂ was suggested by Pickers. (2016), theoretically it could be prone to less bias than the aforementioned tracers due to less uncertainty in the sources and sinks. In particular, it is unaffected by biomass burning and solid biofuel combustion, a large flaw of the CO method.

Universal to all of the methods is the need for the identification of the background molar fractions of said tracers. Background molar fractions are representative of a well-mixed atmosphere where there is no contamination from the local biosphere in the form of respiration or photosynthesis, nor is there any influence from anthropogenic activity (Ramonet and Monfray, 1996). Various methods can be used to produce the baselines which represent the background molar fractions, uncertainties are associated with each. The use of statistical and meteorological baselines will be assessed as part of this study.

1.7. Environmental controls on the activity of the biosphere

The main focus of this study is to understand what causes the variation between days of the diurnal bioCO₂ cycle. Ecosystem scale studies conducted in a whole host of different environments that CO₂ drawdown is primarily driven by solar radiation, whilst relative humidity (RH), air temperature (AT), soil moisture (SM) and nutrient availability also can have significant regulating effects (Guan, *et al.*, 2005, Urbanski, *et al.*, 2007, Vourlitis, *et al.*, 2004). The respiration flux of CO₂ is generally controlled by soil temperature (ST), with warmer temperatures generally creating a better environment for soil borne decomposers (Guan, *et al.*, 2005, Waddington, *et al.*, 2001); SM and AT have also been seen to have significant effect on ecosystem respiration (Price and Black, 1990, Urbanski, *et al.*, 2007). It is clear that the control which environmental variables exhibit over CO₂ fluxes are highly dependent upon many factors such as the type of ecosystem, season and prevailing weather conditions, and as such it is a very large area of research (Baldocchi, *et al.*, 2001, Guan, *et al.*, 2005).

1.8. Vertical mixing

Vertical mixing can be assumed to be related directly to the depth of the planetary boundary layer (PBL) (Denning, *et al.*, 1996), the layer of air directly above the Earth's surface within which the effects of the surface are felt directly (Hennemuth and Lammert, 2006). The PBL's physical characteristics change spatially and temporally as a function of orography, surface cover, season, daytime/night-time and weather. Its depth is primarily determined by the altitude which sensible heat flux is transported, making it a function of solar radiation (Hennemuth and

Lammert, 2006). As shown in Figure 3, a whole host of meteorological conditions interact to determine the PBL depth. Importantly, at the top of the PBL temperature inversions often persist which trap gases in the boundary layer, allowing only very small amounts of transport of gases to

and from free atmosphere above (Bakwin, *et al.*, 1998).

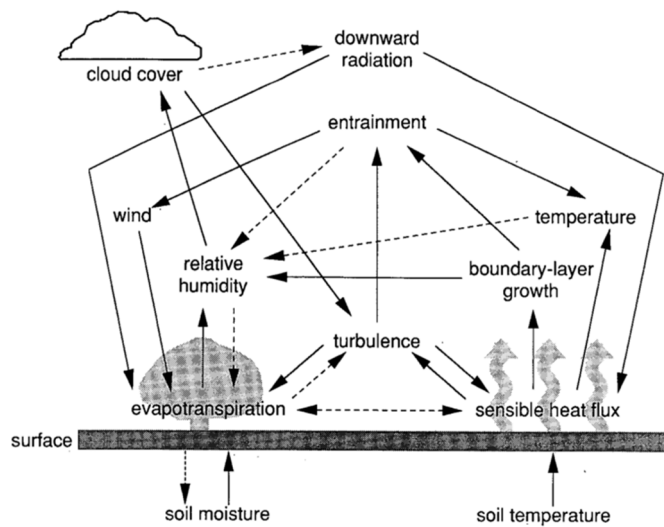


Figure 3 – The interactions between the surface and boundary layer during daytime heating. Positive feedbacks are represented by solid arrows whilst dashed arrows represent negative feedbacks. (Ek and Mahrt, 1994, fig.1).

Under stable conditions the PBL is shallow with a temperature inversion (capped inversion) that traps trace gases and pollutants close to the ground, causing a noticeable increase in the CO₂ (and other gases) mixing ratio (Stephens, *et al.*, 1998). On the other hand when the Earth's surface is warmer than the air above the PBL becomes deep and unstable, a state of free convection is achieved (Hennemuth and Lammert,

2006). In the unstable boundary layer strong turbulent mixing by cumulus entrainment and thermal up/downdrafts act to carry pollutants aloft and mix them throughout the vertical column, effectively diluting the CO₂ signal (Denning, *et al.*, 1996). The vertical mixing is so effective that column studies have shown the CO₂ molar ratio to be homogeneous throughout the air column during unstable conditions (Bakwin, *et al.*, 1998).

During the growing season diurnal variations in the PBL covary with terrestrial biosphere activity. The stable shallow layer occurs at night-time simultaneously to net CO₂ efflux from the terrestrial biosphere, whilst the turbulent deep layer forms during the day when net CO₂ drawdown happens (Figure 4) (Denning, *et al.*, 1996). Both the vertical mixing and terrestrial biosphere activity therefore act to reinforce each other in what has been coined the 'rectifier effect' (Denning, *et al.*, 1995).

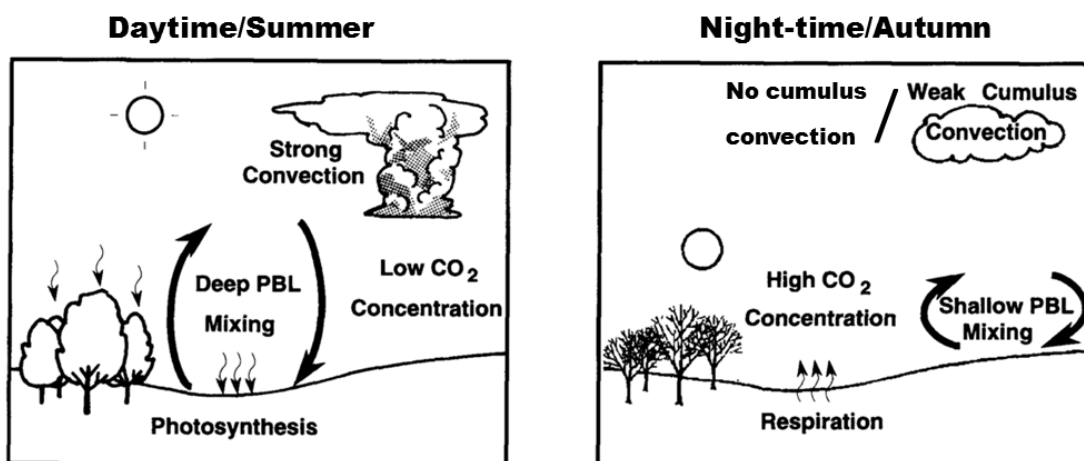


Figure 4 – The vertical CO₂ rectifier effect (Denning, *et al.*, 1996, fig.16) .

1.9. Atmospheric transport

Whilst local conditions impact the bioCO₂ signal, synoptic scale travel of air masses can greatly determine the composition of air masses at a receptor site, a study at Bermuda, Charlottesville for example found that up to 30 % of the atmospheric chemical variability is directly relates to atmospheric transport (Moody, *et al.*, 1989). In general, short-range transport sees the composition affected by point emission sources whilst long range transport sees mixing, physical losses and chemistry affect the composition (Fleming, *et al.*, 2012).

1.10. Aims and objectives

Continuous CO₂, O₂ and CO data from Mace Head atmospheric research station (MHD), Ireland have been used to quantify ffCO₂ mixing ratios using CO and APO. This was carried out over a 18 month period spanning from February 2014 to July 2015. The results from both methods were compared, and the uncertainties in each quantified in an attempt to deliberate the feasibility of APO as an ffCO₂ tracer. An in depth analysis of the causes of uncertainties was performed, with a focus on the baseline generation and emission ratios.

Using the ffCO₂ data the contribution of fossil fuel combustion and the terrestrial biosphere to seasonal and diurnal CO₂ variability was identified. The variability of the diurnal cycle of CO₂ was analysed in depth over an 18 day period in September 2014. Various modelled and measured meteorological variables and their effects on local vegetation were attributed to the diurnal cycle of bioCO₂, whilst the PBL depth and its associated vertical mixing were also taken into account when identifying the causes of short term variation of bioCO₂ molar fractions. Air mass history was also related to variations in the bioCO₂ cycle. Wind speed (WS) and wind direction (WD) data was used in conjunction with back trajectories to decipher the controls on

the measured bioCO₂ molar fraction. In particular the back trajectories were important as they trace the pathway followed by an air parcel ‘up wind’ from the receptor site (MHD), taking into account the synoptic flow of the atmosphere (Fleming, *et al.*, 2012). To my knowledge this is the first study which attributes such a wide range of variables, to the diurnal bioCO₂ cycle.

2. Methodology

2.1. The measurement site

The atmospheric measurements were taken at MHD (Figure 5A) which is located in County Galway on the West coast of Ireland (53°20' N, 9°54' W), just 10 m from the coastline. It can be considered rural, the closest city being Galway, which is 88 km to the East (Salisbury, *et al.*, 2002). A five year study by Simmonds *et al.* (1997) saw 60 % of the air arrive from the Atlantic or Arctic, 32 % from Europe and the other 8 % from Southerly latitudes. This predominantly maritime source location of air masses is also demonstrated in Figure 5B over the entire study period. The predominance of maritime air arriving at MHD makes it an invaluable site for background measurements because oceanic signals are unaffected by anthropogenic/terrestrial signals (Bousquet, *et al.*, 1996). Conversely, its position also allows the study of polluted air masses arriving from continental Europe, Ireland and the British Isles (Salisbury, *et al.*, 2002).

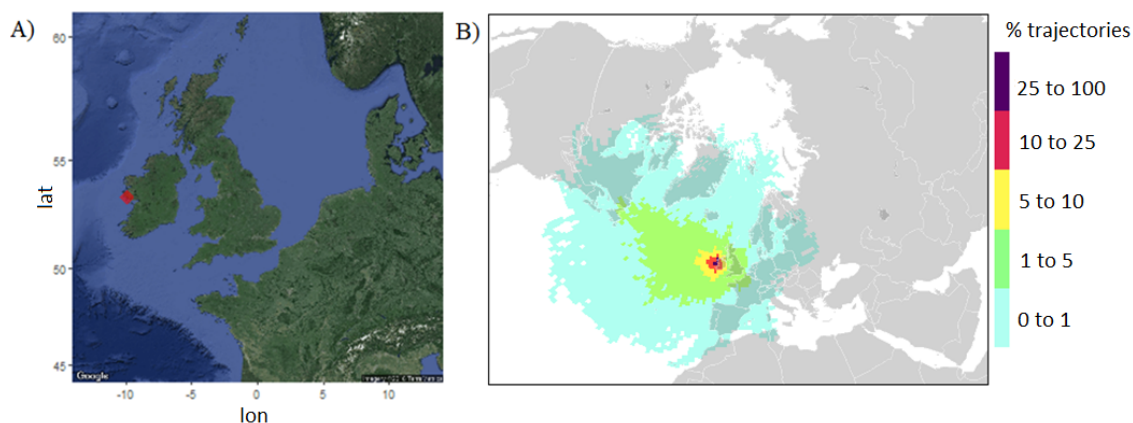


Figure 5 – A) The location of MHD, obtained from: <https://maps.google.co.uk/>. B) Gridded trajectory data from the model HYSPLIT shows the frequency of back trajectories from MHD (2014-2015).

CO₂ and O₂ measurements were collected over an 18 month period from February 2014 until July 2015. A schematic of the system which was used to obtain and measure the air samples is shown in Figure 7; this system is near identical to that described by Wilson (2012). Air is drawn into the system at a height of 25 m as facilitated by two aspirated inlets which act to protect the air from solar radiation, reducing any unwanted solar fractionation (Blaine, *et al.*, 2005). Two inlets are used as a quality control procedure, yet at any one time only air from one inlet is being sampled as selected for by a pneumatically-activated 4-way-valve (V3 in Figure 7), air from the

other inlet is purged.

Moisture from the air is removed prior to sampling by two refrigerator traps in each inlet line (DF1, 2, 3 & 4 in Figure 6), and two in-series chiller traps (DC1 and DC2 in Figure 6). CO₂ mole fraction is determined by the Siemens Ultramat 6E non-dispersive infrared (NDIR) analyser whilst O₂ mole fraction is measured by the Sable Systems Inc. Oxzilla II, both of which are set-up to take measurements every two minutes. Both analysers are in series and therefore receive the exact same air flows. The measurement systems require a flow of the sample gas as well as a stream of gas referenced to the WMO x2007 scale which has constant composition, acquired from the high pressure tanks (WT1 or WT2 in Figure 6). The differential pressure gauge (P22 in Figure 6) maintains zero differential pressure between the sample and reference gas streams. A third chiller trap is also used to ensure the reference gas from WT1 or WT2 is at the same dew point as the sample gas. Regular calibration of the measurement systems are required to maintain accuracy, supplied from calibration gases, the ZT cylinder is used to fix baseline drift. For a more in depth methodology please refer to Wilson (2012).

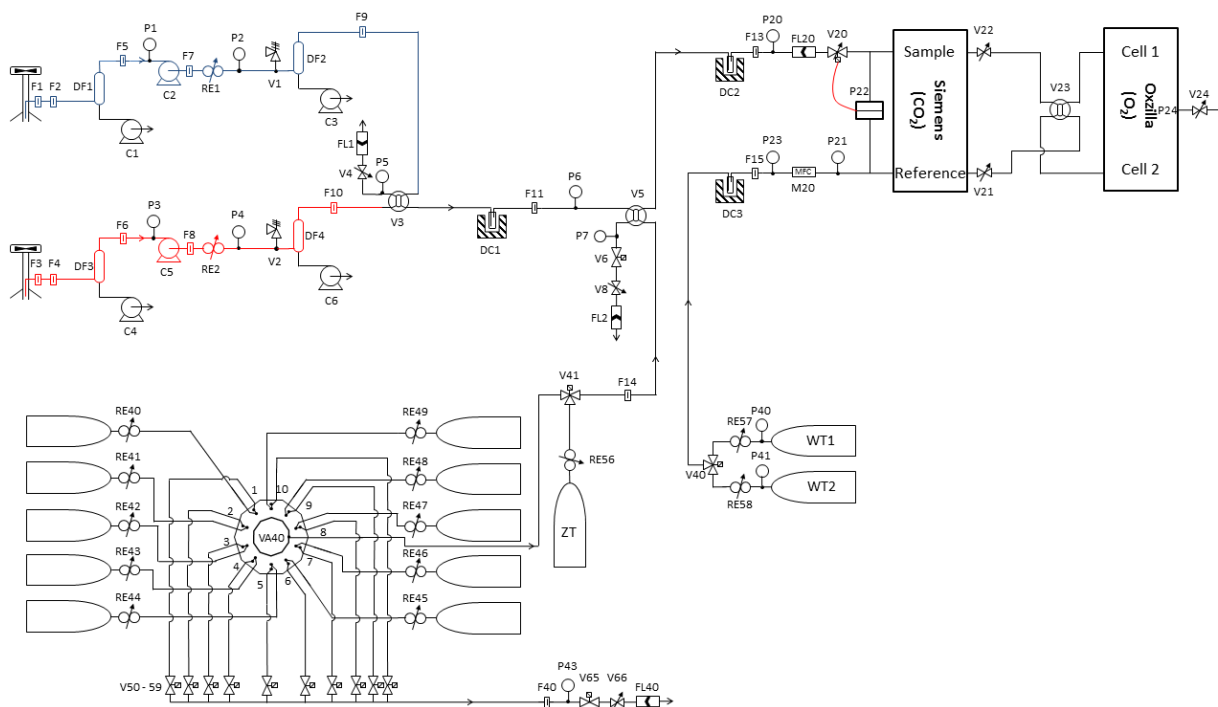


Figure 6 – A schematic of the gas handling system at MHD

2.2. Data handling

CO₂ is recorded in parts per million (ppm), unlike O₂ which is expressed in the more complex units of ‘per meg’. This unit for O₂ is used because its atmospheric content is high and as a result fluctuations in the partial pressure of other atmospheric gases influence the partial pressure of

O₂. To avoid this problem O₂ changes are represented as changes in the ratio of O₂/N₂ as N₂ is far less variable, all of the changes in this ratio are assumed to be due to changes in O₂ (Keeling and Shertz, 1992). The equation for ‘per meg’ is shown in equation 3 as a ratio between the sample gas and a known reference gas (Keeling and Shertz, 1992), for simple conversion 1 ppm $\approx$ 4.77 per meg.

$$\delta(O_2/N_2) = \left(\frac{(O_2/N_2)_{sample}}{(O_2/N_2)_{reference}} - 1 \right) \times 10^6 \quad 3$$

To ensure quality control of the measurements I analysed diagnostic plots which display information about many elements of the handling system such as flow rates, pressures, temperatures and many other variables for the entire study period. I flagged data in the MHD logbook if there was sufficient disparity in the diagnostic plots from what would be expected and/or the raw data was significantly different than would be expected to the point that I knew with certainty it was an artefact of the measurement system. The statistical program R (R Core Development Team, 2008) was used to remove the flagged data points from the raw dataset and simultaneously produce the APO data, using a script written by Penelope Pickers and modified by myself. Alison Craggs performed this diagnostic analysis up until February 2015; I performed it for the rest of the data whilst also making amendments to some of the previous month’s data. APO was generated by using equation 4; where 0.2095 is the standard atmospheric mole fraction of O₂, used to convert CO₂ from ppm to ‘per meg’, 350 is an arbitrary reference, and 1.1 represents the O₂:CO₂ exchange ratio of the terrestrial biosphere (Stephens, *et al.*, 1998).

$$APO = \delta(O_2/N_2) + (1.1/0.2095) \times ([CO_2 - 350]) \quad 4$$

CO data (ppb) for the entire study period was provided by O’Doherty (2015) with a 40 minute resolution. Hourly meteorological data for the entire study period on WS, WD, AT and RH were provided by Ramonet. (2015). Meteorological data in the form of downward short wave radiation flux (radiation), precipitation, cloud cover (CC) and planetary boundary layer (PBL) depth with a three hour resolution were obtained from the Global Data Assimilation System (GDAS) model archive from the National Centre for Environmental Prediction (NCEP) global weather forecast model at <http://ready.arl.noaa.gov/READYamet.php>. GDAS modelled data is provided 24 hours at a time, making it a laborious task to obtain; therefore only data for the September study period was used. Modifying a standard interpolation script which had been produced by Penelope I interpolated the modelled data using linear gap filling to produce a data set with hourly resolution. I also wrote an R script which converted the CO, CO₂ and O₂ data into hourly averages, ensuring matching time stamps, to do this the openair package within R

was required, a package designed for air quality analysis (Carslaw and Ropkins, 2012).

2.3. Baseline generation

Three different baselines spanning the entire study period were generated for the APO and CO data, two being variants of a statistical method, with the other being a meteorological filtering technique (met baseline). Using restricted baseline conditions (RBC) at MHD ‘clean’ maritime air signals were isolated to produce the met baseline; this corresponds to data which was recorded when winds had WD $210 - 290^\circ$ with a $WS > 4 \text{ m s}^{-1}$ and $200 - 210^\circ$ and $290 - 300^\circ$ with a $WS > 8 \text{ m s}^{-1}$ (Bousquet, *et al.*, 1996). Robust extraction baseline signal (REBS) was the statistical technique which was used to generate a ‘flexible’ and ‘stiff’ baseline for both species. REBS was developed by Ruckstuhl *et al.* (2012), and it is a non-parametric technique which assumes that background signals vary slowly compared to the more polluted regional sources and measurements errors are Gaussian distributed with mean 0. Importantly the baseline follows the shape of long term trends and can be used flexibly between measurement sites, whilst it can also handle gaps in the data successfully, an important attribute for long term atmospheric data records in particular. The REBS algorithm is built in to the R package ‘IDPmisc’, as allowing it to be used through R (Ruckstuhl and Locher, 2012).

All three baselines were produced in R, using a template script written by Penelope which I completed and modified to use on the CO and APO data. The ‘span’ variable within REBS specifies the fraction of data points used to compute the fitted values and as such determines the flexibility of the baseline. The flexible baseline was chosen to take into account events on a relatively short-intermediate time scale (days – weeks), whilst not hugging the bulk of the data to the extent that local events influence the baseline. The stiff baseline was determined to vary more steadily with the seasonal variation, excluding most short-intermediate term events. I chose span values of 0.02 and 0.15 for the flexible and stiff baseline respectively. When using REBS algorithm, the baseline is automatically fitted along the bottom of the bulk of the data as local emissions generally raise atmospheric concentrations above the baseline; for APO however fossil fuel burning causes negative excursions in the data and as such the baseline was fitted along the top. To produce this baseline I modified the script to negate the APO data before applying the REBS algorithm, negating the produced baseline resulted in a baseline which fitted along the top of the APO data. Both the REBS algorithm and the meteorological filtering methodology produce baselines which have less data points than the CO and APO data sets, I therefore interpolated the baselines using linear gap filling to ensure a baseline data point for each time stamp of the APO and CO data. This again required a modification of Penelope’s interpolation

script. Only ffCO₂ results produced using the stiff statistical baseline were analysed, a full justification of this choice is contained in the results and discussion section.

2.4. Generating the emissions ratios

The emissions ratio of air at any receptor site changes as a function of the air mass history (Lopez, *et al.*, 2013); emissions databases were therefore used to generate ratios which vary temporally. The R_{CO} values were obtained from the Emissions Database for Global Atmospheric Research (EDGAR) whilst the CO₂ release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) inventory was used to obtain APO:CO₂ emissions ratio (R_{APO}) (EC-JRC/PBL, 2008, Steinbach, *et al.*, 2011). EDGAR is a bottom-up inventory which has a spatial resolution of 0.1 ° x 0.1 °; data on emissions is compiled from a whole range of sources such as the IEA and international road federation (IRF), and it includes information on energy usage, industrial processes, agriculture, land use change and forestry, waste and other anthropogenic sources. The COFFEE dataset has a 1 ° x 1 ° resolution and is produced using the CO₂ emissions from EDGAR with country level information on oxidative rations from the UN energy statistics database, which is based on fossil fuel consumption data (Steinbach, *et al.*, 2011). The most recent edition of both inventories were used: EDGAR 2008 and COFFEE 2014.

Penelope wrote an R script, which I modified to work on the separate EDGAR and COFFEE datasets. It overlaid 96 hour back trajectories on gridded emission ratio data, with back trajectories for every three hours being used for the entire data set. Ratios along the spatial course of each trajectory were averaged to generate the emissions ratio which varied temporally with a resolution of three hours in accordance with the air mass history.

2.5. Calculating the ffCO₂ and bioCO₂ mixing ratios

The ffCO₂ values were produced using equation 5 and 6 for CO and APO respectively, where APO_{bg}/CO_{bg} are the background concentrations as denoted by the baselines and R_{APO}/R_{CO} are the emissions ratios. To do this I generated a script in R which calculated the ffCO₂ molar fraction for each time stamp of the data sets. As can be seen, the APO equation has been derived directly from the widely used CO method by substituting in the APO values for the corresponding CO values.

$$ffCO_2 = \frac{CO - CO_{bg}}{R_{CO}} \quad 5$$

$$ffCO_2 = \frac{APO - APO_{bg}}{R_{APO}} \quad 6$$

Using a simple mass balance equation (equation 7) the bioCO₂ was determined. To do this I produced an R script which selectively emitted any negative ffCO₂ mixing ratios which had been spuriously generated from inaccuracies in either the baseline emissions ratios or measurements. Equation 8 was then applied to each time stamp. It is important to note that the methodology here does not produce the absolute values for the bioCO₂ (abioCO₂) as the molar fraction of the bgCO₂ component has not been removed; equation 8 demonstrates how the absolute bioCO₂ values can be calculated. Leaving the bgCO₂ component in the bioCO₂ data still allows the diurnal variability to be analysed in depth, but removes the uncertainty which would be associated with implementing another baseline, the stiff statistical CO₂ baseline (see appendix, Figure 1).

$$bioCO_2 = obsCO_2 - ffCO_2 \quad 7$$

$$abioCO_2 = obsCO_2 - ffCO_2 - bgCO_2 \quad 8$$

In depth analysis of the data throughout the year showed noise which could not be representative of bioCO₂. Further examination showed that there is too much noise in the O₂ measurements and there also appeared to be a fractional difference in the data results from the two air inlets.

Although this noise is relatively small and has a negligible effect on the seasonal cycle, it is enough to invalidate observations of the bioCO₂ diurnal cycle. The period September 4th – 22nd is the only period where this noise is at an acceptable level and a consistent diurnal cycle can be noticed. During this period all of the data is also coming through one line, the red line meaning that there are no uncertainties associated with switching between the two inlets. Unfortunately during the September study period there was no CO data, meaning the two methods could not be compared at short time scales, but the uncertainty analysis was used to quantify the differences between methods which occur over the entire 18 months.

2.6. Uncertainty production

There are three sources of uncertainty in the production of the ffCO₂ data: measurements, ratio and baseline uncertainty. The standard deviation (σ) was used to quantify the uncertainty in each of the three sources; quantifying the dispersion of the data from its mean. Showing this as a percentage of the uncertainty allowed an easy comparison between the different sources of uncertainty. Summing in quadrature could be used for a more accurate way to combine the three sources into one uncertainty for the results, yet it is too extensive a task for this particular project, furthermore it is highly likely that the largest uncertainty will dominate the total uncertainty, so it was also deemed gratuitous (Pickers, 2016, personal communication).

The data sets were split up into five periods by season so that (i) much of the annual variation in the data (seasonal cycle) to be removed, as this natural seasonal cycle would add to the uncertainty and (ii) the short term fluctuations of CO₂ are different per season. I chose the periods Spring (March, April and May), Summer (June, July and August) and Autumn (September, October and November) for the 2014 data as well as Winter (December, January and February) and Spring for the 2015 data. Stretches of uninterrupted data were analysed for each period to avoid including any step changes which might occur as a result of calibration or forced pauses in the measurement process. Each stretch of data was 340 points in length as this was the longest length of data which could be applied to all five periods without significant interruption. The selected data points were located as close to the middle of each season as was possible, starting on: 28th March, 1st July, 18th October, 7th January and 8th April.

To generate the baseline uncertainty I took the σ of the absolute differences between the ffCO₂ results produced by the met baseline and the stiff statistical baseline (Equation 9). The difference between these two baselines was chosen because they cover the upper and lower limits of values produced by the baselines (all three of which are legitimate techniques), leading to the largest and smallest ffCO₂ values. When quantifying the ratio and measurement data the standard deviation of the absolute values were used (Equation 10 and 11), these analyses were applied to the APO and CO data.

$$\text{Baseline Uncertainty} = \frac{\sigma(|\text{ffCO}_2(\text{stiff baseline}) - \text{ffCO}_2(\text{met baseline})|)}{\text{ffCO}_2(\text{stiff baseline})} \quad 9$$

$$\text{Measurement uncertainty} = \frac{\sigma(\text{Measurement})}{\text{Measurement}} \times 100 \quad 10$$

$$\text{Ratio uncertainty} = \frac{\sigma(\text{Ratio})}{\text{Ratio}} \times 100 \quad 11$$

2.7. Data analysis

As the diurnal cycles of both the terrestrial biosphere and PBL are strongly determined by solar forcings the data was split up into daylight and night-time hours; to do this I wrote an R script which used openair to split the data. The split was generated using the geographical location of MHD, date, time and astronomical algorithms which estimate the position of the sun relative to the location of MHD according to a National Oceanic and Atmospheric Administration (NOAA) method (Carslaw and Ropkins, 2012). Cross correlation tests were performed on the daylight-split data in R using a script which I wrote; this allowed any lag to be identified between the meteorological variables and bioCO₂ activity. The openair package in R was used to generate

‘correlation plots’ which allow for easy visualisation of the correlations between variables. The openair package also provided a good platform to display time-series data graphically.

2.8. Determining the air mass history

WS and WD regimes were used to identify the source of air masses reaching MHD; in particular this data was most helpful at identifying signals which may have a local source. Land and sea breezes are prominent at the coastline as differential heating of the land compared to the ocean generates mesoscale winds which flow landwards during the day and seawards during the night, but at higher altitudes (Bouchlaghem, *et al.*, 2007). Breezes are often distinctly different from the general synoptic air flow, as such previous studies at MHD (i.e.(Cape, *et al.*, 2000, Salisbury, *et al.*, 2002)) have removed these from the data by assuming the breezes have a $WS < 3 \text{ ms}^{-1}$.

Under breeze conditions it is highly likely that the bioCO₂ values have a local influence (Cape, *et al.*, 2000), as such I wrote an R script to isolate these periods. A conditional probability function plot (CPF) was generated to demonstrate the probability of source locations which are most likely to be upwind when the bioCO₂ measurements in the 95th percentile or above are recorded (Fleming, *et al.*, 2012). The CPF plot was generated using equation 12 where m_θ is the number of samples in wind sector θ with the $\geq 95^{\text{th}}$ percentile molar fractions, and n_θ is the total number of samples; again this was plotted using functionality within openair.

$$CPF = m_\theta / n_\theta \quad 12$$

Hybrid Single Particle Lagrangian Integrated Trajectory model (HYSPLIT) was used to generate the 96 hour back trajectories for the September study period (Draxler and Hess, 1998). A model which was originally designed to help the Bureau of Meteorology respond to atmospheric emergencies. It uses a hybrid between Lagrangian and Eulerian modelling; the former uses a fixed dimensional reference system whilst the latter uses a reference system which follows the trajectory through time and space (Fleming, *et al.*, 2012).

Openair has the function to import pre-calculated HYSPLIT back trajectories, a pre-requisite for the HYSPLIT model is gridded field meteorological data, the Global NOAA-NCEP/NCAR reanalysis data archives were used to generate the HYSPLIT back trajectories, as suggested by the HYSPLIT developers (Kalnay, *et al.*, 1996). Although not the meteorological field data with the highest resolution (2.5 °), it was used because it is guaranteed to be practicable over a long time period, and furthermore is more widely available than other data sets (Draxler and Hess, 1998). Back trajectories are run at 3-hourly time intervals, starting at a height of 10 m, propagating 96 hours back in time. I wrote a script to import these trajectories, and merge the

trajectory data with bioCO₂ concentration data; trajectories were then colour coded according to the bioCO₂ concentration at MHD using the jet colour scheme (Appendix, Figure 3).

Whilst the trajectories alone can be helpful, they are not easy to visually analyse. Cluster analysis was performed to group back trajectories based on their geographical origin. An angle distance matrix was used to decide on the similarity/dissimilarity between trajectories, alternatively, Euclidean distance matrix analysis could have been used. Carslaw. (2012) suggest that the trajectory separation method be chosen based on post-processing data analysis. The angle distance matrix method explained the varying characteristics of the bioCO₂ diurnal cycle more coherently and was therefore chosen. Clusters were generated by using equation 13.

$$d_{1,2} = \frac{1}{n} \sum_{i=1}^n \cos^{-1} \left(0.5 \frac{A_i + B_i + C_i}{\sqrt{A_i B_i}} \right) \quad 13$$

Where

$$A_i = (X_1(i) - X_0)^2 + (Y_1(i) - Y_0)^2 \quad 1$$

$$B_i = (X_2(i) - X_0)^2 + (Y_2(i) - Y_0)^2 \quad 2$$

$$C_i = (X_2(i) - X_1(i))^2 + (Y_2(i) - Y_1(i))^2 \quad 3$$

and X₀ and Y₀ are the co-ordinates of MHD.

One thing that back trajectories lack is the ability to attribute carbon signals to a source location, a concentration weighted trajectory (CWT) plot was therefore produced to determine the source and sink locations of bioCO₂ (Cheng, *et al.*, 2013). CWT analysis acts to attribute the bioCO₂ molar ratio at MHD to all segments of the corresponding trajectory with equal weighting as in equation 14. This again is another function of openair.

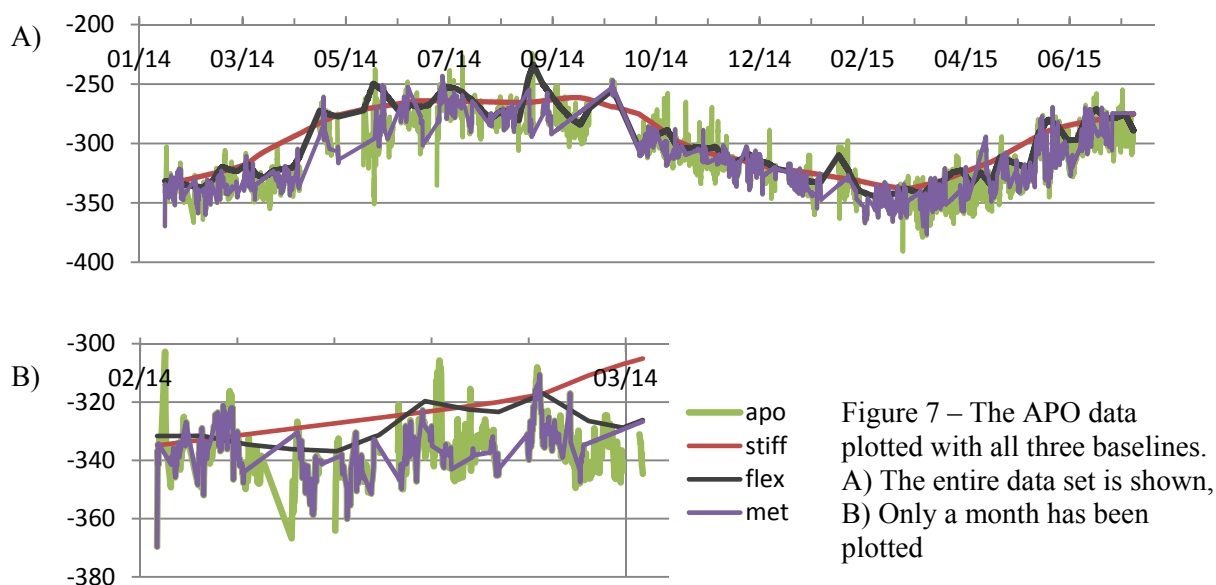
$$\ln(\overline{C_{ij}}) = \frac{1}{\sum_{k=1}^N \tau_{ijk}} \sum_{k=1}^N \ln(C_k) \tau_{ijk} \quad 14$$

3. Results and Discussion

3.1. Baseline selection

In Figure 7 the three baselines are demonstrated on the APO data set. Plot A demonstrates how each of the baselines fit the entire study period. Importantly, all three baselines strongly represent the data's seasonal cycle, which is important as the background air varies with global sinks and sources (Levin, *et al.*, 2008). Plot B shows how the baselines vary on a short-intermediate time period (mid-February to mid-March was chosen arbitrarily). A student's t-test

suggests that all three baselines are statistically different from each other ($P < 0.01$) even though all three methods are legitimate techniques used to represent the background air at MHD.



There are various issues with the met baseline, the most recognisable characteristic being that it varies dramatically over very short time scales; on 12th February for example it sees a 25 per meg drop in just a three hour period, and that is common place throughout the study period. As the background air is affected by global scale sinks and sources it would be expected to see seasonality, not so much short-term variation (Keeling, *et al.*, 1995). The vast majority of large short term variation can be attributed to polluted sources (terrestrial, anthropogenic or other). The large variability in the met baseline would therefore lead one to believe it is capturing some polluted signals, whilst its placement in the middle of the bulk of APO and CO (see appendix, Figure 1) data with an average value of -309.8 per meg (for APO) also raises questions about its legitimacy. As aforementioned anthropogenic fossil fuel combustion causes negative excursions of the APO data, its position near the middle of the APO data set means that a vast amount of positive excursions are included which lead to negative ffCO₂ values, an impossibility. Furthermore, this implies that the bgCO₂ is under influence of fossil fuel combustion more than the local data, again this seems unlikely.

Perhaps these issues are caused as WS and WD are often not representative of the wider synoptic flow in the atmosphere (Fleming, *et al.*, 2012). Data assumed to be from the Atlantic Ocean because it was isolated by the RBC may well have originated from inland where there are many anthropogenic and terrestrial sources. Another contention can be seen as the meteorological filtering cannot be applied to all measurement sites, making comparisons between data sets difficult (Ramonet and Monfray, 1996).

Statistical baselines are advantageous as they can be used homogeneously at different measurement sites, universal use of this techniques would lead to easier comparisons of background data (Ruckstuhl, *et al.*, 2012). Importantly both of the statistical baselines are fitted along the top of the bulk of the APO data and along the bottom of the bulk of the CO data as expected, with average values of -299.7 and -296.9 per meg for the flexible and stiff APO baselines respectively. This means that compared to the met baseline the data sees fewer negative excursions from the baseline for CO and relatively fewer positive excursions in the case of the APO data. It was the stiff baseline which in fact saw the fewest unexpected excursions and as a result it produced the least amount of negative ffCO₂ results, all results of which must be erroneous, providing further impetus to choose the stiff statistical baseline. Judging the span variable and subsequently the baseline flexibility so that it is accurate of background air was one of the biggest challenges, especially as the scientific literature contains little information about generating the baselines and the behaviour of background air. To err on the side of caution this also led to me choose the stiff baseline as it removes far more of the short-intermediate term variability which may be caused by anthropogenic/terrestrial events. One thing which must be considered is that the stiff baseline may well miss out on some short-term background fluctuations, and as such some fluctuations in the generated ffCO₂ or bioCO₂ data could be caused by variations in background air. On top of this, it is not clear whether any statistical baseline techniques can effectively remove low levels of polluted air (Ryall, *et al.*, 2001).

3.2. Uncertainty analysis

The data from the uncertainty analysis in Table 1 shows that as a tracer APO has less uncertainty than when using CO to quantify the atmospheric mixing ratio of ffCO₂. As part of a simplification within this report the largest uncertainty associated with each method is assumed to dominate the total uncertainty; the APO uncertainty is therefore far less than the CO uncertainty, 49.6 and 221.1 % averages respectively. Whilst the baseline uncertainty dominated the ffCO₂(APO) it was the uncertainty in the emissions ratio which dominated the ffCO₂(CO) uncertainty. In fact for all three of the sources of uncertainty the uncertainty was larger for CO than APO. It is also clear that the uncertainties vary dramatically between the seasons, being higher in Spring and Summer for both sets of measurements, the CO baseline and R_{APO}, no such trend was seen for the APO baseline and R_{CO}.

Table 1 – Uncertainty values for entire study period, as APO is a function of CO₂ and O₂ the uncertainties for the separate measurements are displayed in the table.

	Baseline (%)		Emissions ratio (%)		Measurement (%)	
	APO	CO	APO	CO	CO ₂ /O ₂	CO
Spring 2014	17.0	76.6	16.2	182.9	0.74/3.12	6.42
Summer 2014	85.2	116.7	16.2	110.0	1.67/7.38	13.11
Autumn 2014	45.1	91.0	15.3	261.6	0.55/2.50	8.22
Winter 2014/2015	50.2	56.8	10.0	303.6	0.26/1.91	7.00
Spring 2015	50.2	72.7	13.0	247.4	0.98/3.51	10.90
Average	49.6	82.8	14.14	221.1	0.84/3.68	9.13

The large amount of uncertainty in the CO measurements may be a result of large spikes in the molar fraction which occur quite regularly (Appendix, Figure 2). The O₂ and CO₂ data shows far less variance, and as such have less of an uncertainty. Compared to the other sources of uncertainties, the measurement uncertainties are negligible. I do however propose that the variation of the measurement data and baseline data are linked, and this is why the baseline uncertainty is far larger for CO. As the APO and CO baselines were produced using the same methods it may be possible that the large fluctuations in CO measurements acted to increase baseline uncertainty. If the large spikes in the CO data are a result of error in the sampling methodology the baseline uncertainty could decrease under revised conditions. Using this data alone it is not definitively possible to assume that the CO baseline will always be more certain than the APO baseline. On the other hand it is always possible that the CO molar fraction varies more in the atmosphere, potentially due to the larger amount of sources and sinks than APO.

The emissions ratio produces another large source of uncertainty, for CO this is unacceptably high (221.1 % average). This is in agreement with Vardag *et al.* (2015) who found R_{CO} was the greatest limitation when using CO. Sporadic elevations in the R_{CO} data occur, many of which must be erroneously included in the inventory. On 90 occasions R_{CO} exceeds 50 ppb ppm⁻¹, occurring on 29 different days. These R_{CO} spikes are far larger than the median R_{CO} of 1.46 ppb ppm⁻¹, furthermore in Heidelberg weekly variations of R_{CO} were just ±5.6 ppb ppm⁻¹, far less than these seen at MHD, further suggesting the data is erroneous (Vogel, *et al.*, 2010).

Whilst these spikes certainly appear erroneous, Lopez *et al.* (2013) showed that EDGAR 4.2 R_{CO} is lower than both the CITEPA (AirParif and national emission inventory by the Centre Interprofessionnel Technique d'Etude de la Pollution Atmosphérique) and IER (the institute for energy economics and the rational use of energy inventory) emissions ratios for Paris. They also

found that the R_{CO} ratio was less than what they determined by using a top down atmospheric approach. Even though these results are from a different geographical location it is plausible that R_{CO} values are also too low for the air reaching MHD. In fact as to be mentioned later the $ffCO_2(CO)$ values are quite a bit higher than would be expected at such a rural location, which may well be caused by R_{CO} values being too low. Gamnitzer *et al.* (2006) also acknowledged the significant differences between EDGAR and IER, much of which are due to the different sources on fossil fuel usage. In particular EDGAR does not correctly include diurnal cycles of emissions. The EDGAR data sees the inclusion of various non-fossil sources, which again may add to the R_{CO} uncertainty; Gamnitzer *et al.* (2006) determined that removing non-fossil fuel sources could decrease R_{CO} by as much as 17 %, having a significant effect on $ffCO_2$ quantification, perhaps this could be corrected for in the future to increase the reliability of $ffCO_2(CO)$ results.

The uncertainties are far more acceptable for the R_{APO} data; it has a range of 0.22 per meg ppm^{-1} compared to the R_{CO} range of 118.53 ppb ppm^{-1} . It is clear that its values are in a more defined and smaller range, meaning $ffCO_2$ values are less affected by variance in R_{APO} . Whilst there is nothing in the scientific literature to my knowledge which analyses the uncertainty of COFFEE, one would expect it to be more constrained as only fossil fuel and liquid/gaseous biofuel combustion affect its short term variation. Obviously these emissions ratios are prone to uncertainties in the same way as bottom-up fossil fuel emissions inventories, but what makes use of these emissions ratios more reliable is that they are constrained within defined bounds as the emissions ratios R_{APO} and R_{CO} are well-known (Keeling, 1988, Vardag, *et al.*, 2015)(Pickers, personal communication).

Whilst there are almost certainly issues with the inventories, the method with which the ratios were extracted may have also contributed to the uncertainties. HYSPLIT trajectories were used to generate temporally varying ratios. The problem with back trajectories is that these only represent one route which the air parcels could have taken, in reality vertical and horizontal mixing will lead to a wide range of source locations leading to MHD at any one time (Ryall, *et al.*, 2001). Trajectories also struggle to take into account land and sea breezes, potentially inducing false values (Ryall, *et al.*, 2001). This method also assumed that each section of the trajectory provided an equal amount of the ratio to the received air mass, an improbability. On top of this, to make the process computationally viable the spatial span of the COFFEE and EDGAR data had to be reduced, on occasions the trajectories spanned outside area of the inventory's region and as such R_{APO} along the entire trajectory could not be calculated. It is also important to remember that COFFEE was from 2014 and EDGAR from 2008 and as such don't match the data points temporally, potentially making the ratios significantly incorrect.

3.3. CO₂ seasonal variability

As was highly likely to occur there are large differences between the CO and APO derived ffCO₂ results (Figure 8), the most obvious difference being that the CO results were higher with a mean of 11.41 ppm compared to 8.14 ppm for ffCO₂(APO). The ffCO₂(APO) data monthly means ranged from 2.6 ppm in December 2014 to 16.08 ppm in April 2014 (an especially polluted month), the next most polluted month was May 2015 with 12.2 ppm. The ffCO₂(CO) data saw ranges from 2.6 ppm in June 2014 to 28.44 ppm in March 2015. In comparison, the similarly rural site, Lutjewad, Netherlands sees monthly mean ffCO₂ of 3 – 10 ppm, the MHD data therefore appears to be very high, especially ffCO₂(CO) (Levin and Karstens, 2007). In fact the highly populated region of Heidelberg, Germany shows mean monthly ffCO₂ molar fractions ranging from 5 – 20 ppm, values which are not too dissimilar from the APO values., The ffCO₂(CO) data saw monthly means even larger than that of Heidelberg, further suggesting that the R_{CO} was far too low on average (Levin, *et al.*, 2003). There is no correlation ($r = 0.12$) between the two ffCO₂ data sets, and as such the two results could be considered strongly different from each other. Furthermore, negative spikes which occurred in equal quantities for ffCO₂(APO) and ffCO₂(CO) have been removed which must primarily be artefacts of the baseline or erroneous measurements and ratios. Spikes also occurred in the positive direction, reaching over 50 ppm for the APO data, and over 150 ppm for the CO data; many of these would be considered erroneous, especially as Lopez *et al.* (2013) recorded maximum ffCO₂ results of 41.9 ppm in Paris in February 2013, a far more urban area then MHD.

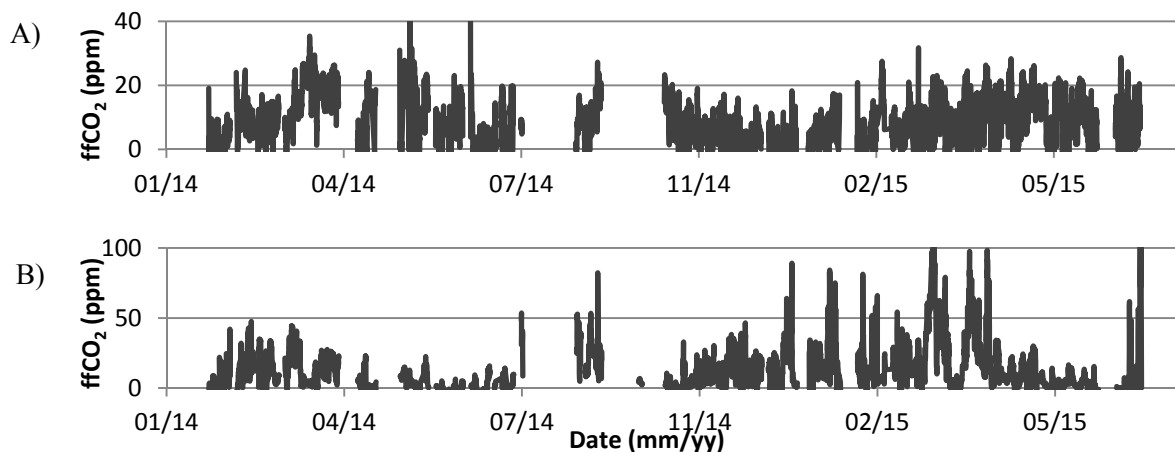


Figure 8 – The ffCO₂ values as determined by using (A) APO as a proxy and (B) CO as the proxy. The y-axis have been reduced for clarity, cutting out some of the ffCO₂ spikes.

Even though there are large differences between months in the ffCO₂ values, this data is not strong enough to imply that ffCO₂ has a strong overall influence on obsCO₂ seasonality. The APO data suggests a minimum in late Autumn and Early winter followed by a peak in mid

Spring, the CO data sees the peak in early Spring and the minimum in early Summer. It is clear that the terrestrial biosphere (and oceanic influence as it is included in the baseline) contributes mostly to the seasonal cycle of obsCO₂, with an R² of 0.41 (Figure 10), the distinct similarities between obCO₂ and bioCO₂ can be seen in the appendix, Figure 2.

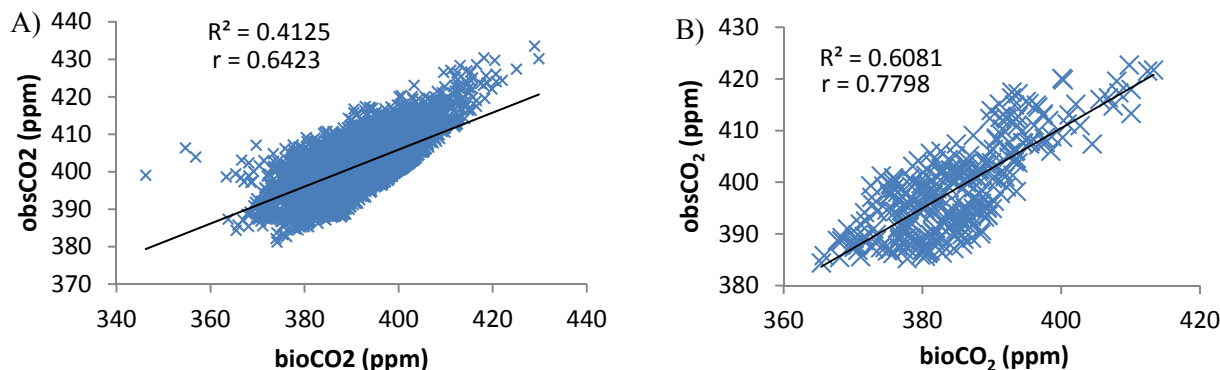


Figure 9 – A linear regression relating obsCO₂ to bioCO₂ for A) the annual variability and B) the diurnal variability.

3.4. CO₂ diurnal variability

Unlike large urban areas such as Paris and London which have both been shown to have relatively large daily cycles in ffCO₂ (Hernandez-Paniagua, *et al.*, 2015, Lopez, *et al.*, 2013), no distinct ffCO₂ diurnal cycle was seen at MHD, which is almost certainly due to its rural location. Nor was there any difference between weekends (6th–7th, 13th–14th and 20th–21st) and weekdays. What is noticeable however is that the ffCO₂ data can be separated into three distinct sections, the 4th – 8th and 20th – 22nd which see lower average ffCO₂ molar fractions than the 9th – 19th.

As shown in Figure 9, the diurnal variability of obsCO₂ is very largely determined by bioCO₂ (r = 0.78). On average the obsCO₂ and bioCO₂ molar fractions increase overnight, peaking just before 06:00, then decrease during the day, reaching a trough at 18:00 on average (Figure 10). This shape is characteristic of a net uptake of CO₂ by the terrestrial biosphere during the day by net photosynthesis, followed by a release of CO₂ overnight when photosynthesis and plant/soil respiration continue. Over the September study period there is a large amount of variation between days in the amplitude and timing of bioCO₂ uptake/release. On the 4th–5th for example the net bioCO₂ efflux is 36.5 ppm, reaching a maximum of 413.3 ppm; on the 16th however the bioCO₂ peaks at just 390.4 ppm, with a net uptake prior to this of 22.6 ppm. Four distinct sections in the bioCO₂ data can be defined where the characteristics of the diurnal cycle are notably similar; importantly these can be related to the meteorological variables (excluding radiation), air mass history and ffCO₂ values. Section (A) covers 4th – 5th where the bioCO₂ amplitude is large and well defined, (B) 6th – 8th where the bioCO₂ diurnal cycle is small in

amplitude, (C) 9th – 19th (midday) where the diurnal cycles are large in amplitude and easily definable and (D) 19th (midday) – 21st where the diurnal cycles undefinable. It is clear that section C of the bioCO₂ data is linked to the period with the highest average ffCO₂, on the other hand sections B and D which have the least definable bioCO₂ diurnal cycle are linked to periods with the lowest ffCO₂ values. Section A also sees the low ffCO₂ values, but a relatively defined diurnal bioCO₂ cycle.

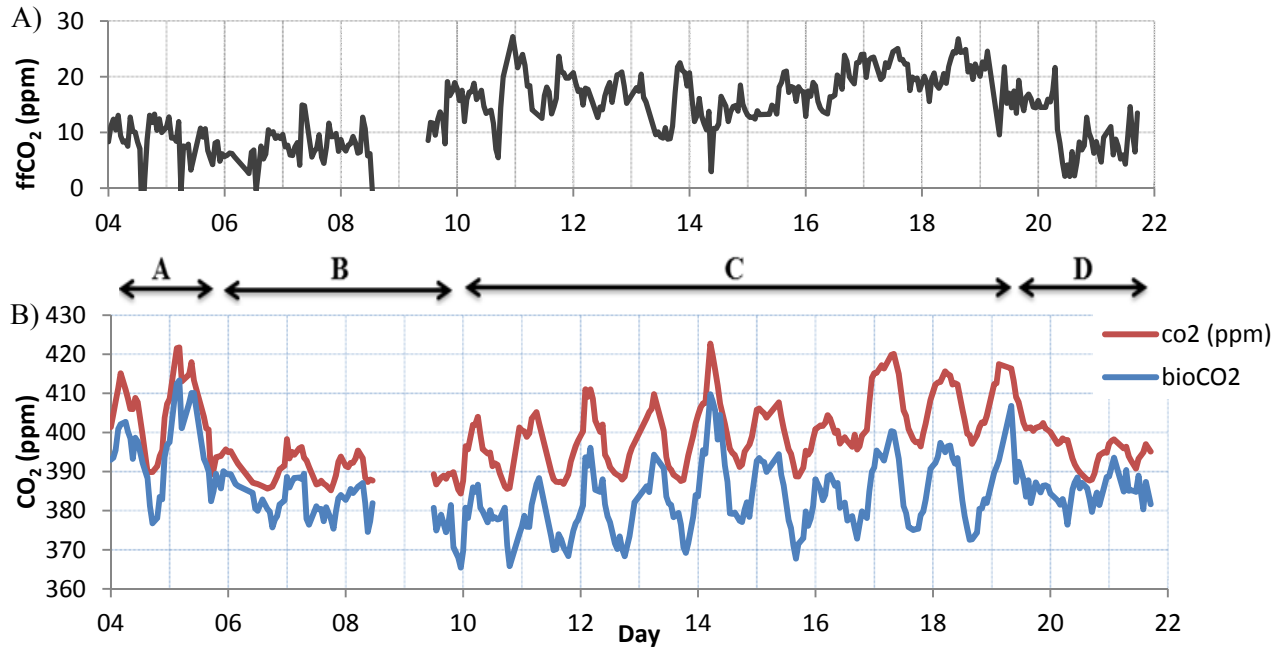


Figure 10 – A) FfCO₂ over the September study period. B) The obsCO₂ and bioCO₂ diurnal cycles over the September study period.

3.5. Meteorological effects on the local biosphere

Moderate correlations were found between bioCO₂ and RH, radiation, PBL depth and AT, with a total correlation of 0.64, -0.60, -0.49 and -0.38 respectively. Being a function of solar forcing all four of these variables exhibited their own diurnal cycle, AT and RH preceding bioCO₂ by one hour whilst PBL depth and radiation precede bioCO₂ by 2-3 hours. Lags were removed before generating the correlation plot in Figure 11. AT, radiation and PBL depth show a stronger correlation with bioCO₂ during daylight than night-time, whilst the AT:bioCO₂ correlation is in fact inexistent at night-time. In particular these variables (excluding PBL depth) would be expected to influence the activity of local vegetation, most of which is temperate grassland and shrubbery, in the past most of the local terrestrial biosphere activity in the past has been attributed to a peatland located adjacent to the measurement site (Derwent, *et al.*, 2002).

To imply causality between the different meteorological variables and bioCO₂ drawdown/release two principles must be observed. (i) The cause must occur before the effect and (ii) the causal

variable must contain special information about the bioCO₂ variation which must only be seen in said variable (Granger, 1988). As the bioCO₂ diurnal cycle lags that of the four variables they are all indicative of a causal relationship and the first principle is fulfilled.

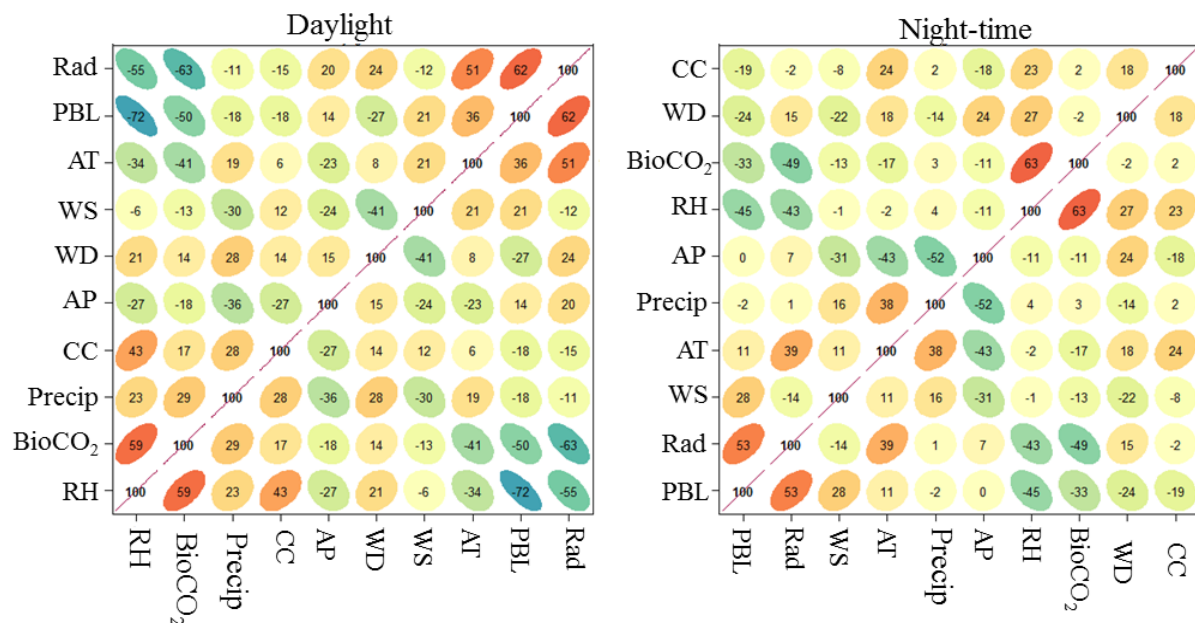


Figure 11 – A correlation plot for the September study period, daylight correlations are shown on the left whilst night-time correlations are illustrated in the right hand panel. The shape, colour and numerical value donate the correlation between said variables.

RH which has the largest total correlation with bioCO₂, seeing a marginally larger correlation during night-time (0.63) than daytime (0.59). Figure 12 shows how RH increases overnight and then decreases throughout the day, reaching a minimum in the afternoon. There is some observable covariation between bioCO₂ and RH, in particular section C which sees the largest and most defined diurnal cycles in RH (peaking at ~90 % and forming a trough at 60-75 %) simultaneously to the large well-defined bioCO₂ diurnal cycles. On the other hand section A sees no RH diurnal cycle at all, even though the bioCO₂ cycle is large; whilst section D which sees no diurnal bioCO₂ cycle sees a significant diurnal cycle in RH. Furthermore it does not seem to provide any individual control over bioCO₂, for example on the 14th the bioCO₂ peaks at ~410 ppm but the RH sees no excessive peaks compared to the surrounding days which could have caused this.

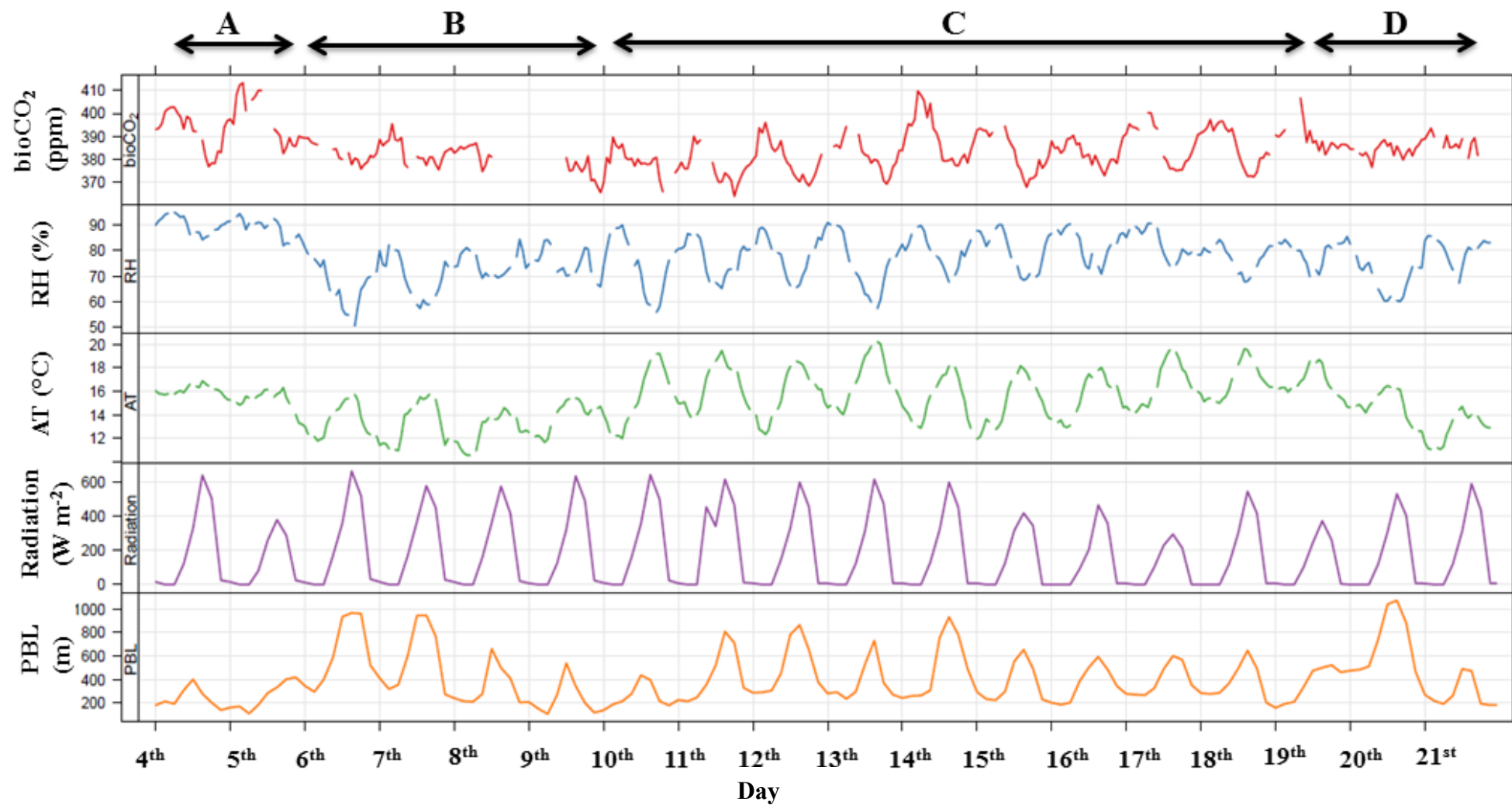


Figure 12 – Plots of bioCO₂, RH, AT, radiation and PBL depth for the September study period, the distinct sections of data are illustrated at the top

It is likely that RH does have some control on terrestrial activity, probably through control of the stomata, pores which are responsible for the exchange of gas between plants and the atmosphere necessary for photosynthesis (Boyer, 2015). Guard cells control the stomata and can act as humidity sensors, shutting the stomata during periods of decreased RH and opening them again when the RH is sufficiently high, effectively controlling photosynthesis (Lange, *et al.*, 1971). This well studied physiological mechanism could explain much of the daylight correlation, especially as net photosynthesis ceases an hour after RH reaches its minimum. The strong correlation at night is also indicative of RH having some sort of control over respiration as well. Soil moisture (SM) can control respiration rates, increasing SM in peatlands has shown to increase respiration and CO₂ efflux, with a peak in respiration occurring when the soil is almost saturated at 92% water content (Kuiper, *et al.*, 2014). In tallgrass prairie ecosystems SM has been shown to determine 26 % of the soil respiration rates (Mielnick and Dugas, 2000). In fact SM has also been shown to have an influence on photosynthesis; peatlands with a low SM content due drought conditions rapidly turn from a net carbon sink to source (Kuiper, *et al.*, 2014). If RH has a significant effect on SM, this could explain its control over respiration; however no definitive link can be made without SM measurements.

AT sees an increase during the day, reaching a peak in the afternoon, before decreasing overnight and seeing a trough in the early hours. There is lots of variation between days in the AT, similarly to section RH, section A sees no diurnal cycle even though the bioCO₂ diurnal cycle is significant, whilst section D sees two defined diurnal AT cycles in the absence of diurnal bioCO₂ cycle. Section B on the other hand sees smaller diurnal cycles in AT than the rather defined diurnal cycles in AT in section C, these two sections suggest that there is some covariation between AT and bioCO₂. AT has in fact been shown to affect the rate of peatland net respiration, with the highest CO₂ production rates occurring at 20 °C, on the 13th-14th in fact when AT reached 20 °C (Waddington, *et al.*, 2001), and this is when the largest respiration flux was seen. Similarly, soil temperature (ST) has been shown to control respiration in tallgrass prairie, accounting for 46 % of respiration CO₂ flux variability (Mielnick and Dugas, 2000). AT is also the main determinant for respiration rate, respiration increasing exponentially with AT (Urbanski, *et al.*, 2007). At MHD however, the lack of AT:bioCO₂ correlation at night-time could suggest that it has relatively little control over respiration at all. During the day the bioCO₂:AT correlation is moderately anti-correlated (-0.41), maybe a marker of its influence on photosynthesis. A study in a boreal forest in Hyytiälä, Finland proved AT to be the limiting factor for photosynthesis during Winter, Spring and Autumn, as such AT fluctuations directly impact the forest's productivity (Wu, *et al.*, 2012). As suggested by the only moderate anti-

correlation the limited influence of AT over photosynthesis, may suggest that it is not a limiting factor, especially as conditions at MHD are warmer than Hyytiälä. Similarly Jia et al. (2015) found that AT (and solar radiation) were determining factors for carbon fluxes in a subtropical forest in winter, when the temperature was lower. It seems highly likely that AT is not a limiting factor for either photosynthesis or respiration and as such has no significant control over either process. As with SM, measurements on ST would be desirable as it could show its own controls over the activity of the terrestrial biosphere, probably respiration.

It is well documented on a plant and ecosystem level that photosynthesis is primarily driven by light intensity, and this can be seen by the (anti)correlation radiation has with bioCO₂ during the day (-0.63), the strongest daytime correlation (Gu, *et al.*, 2003, Guan, *et al.*, 2005, Urbanski, *et al.*, 2007). Every night radiation drops to 0 at 21:00, before increasing at 06:00, reaching a peak at 15:00. It is clear that singular events cannot be attributed to changes in the available radiation, nor can it be separated out into the four pre-defined sections. The correlation however is very strongly indicative of its control over photosynthesis, with higher incident radiation rates forcing higher rates of carbon drawdown. The moderate radiation:bioCO₂ anti-correlation at night-time could be indicative of the cessation of photosynthesis being associated with radiation dropping to 0 Wm⁻². To better understand the effect of the radiation, information of the wavelengths incident upon the surface at MHD will be required, as only light of wavelength 400 – 700 nm lead to photosynthesis (Alados and Alados-Arboledas, 1999).

It is clear that none of the diurnally varying meteorological variables alone can fulfil the premise for causality. Variations in the bioCO₂ diurnal cycle appear to be forced by a summative effect of each of the variables. Furthermore, it is also possible that none of the meteorological variables are limiting factors, and as such their influence on bioCO₂ not witnessed to its full extent (Gu, *et al.*, 2003). Importantly, hysteresis, filtering, time lags, amplification and chaos have all been witnessed in previous ecosystem studies (Baldocchi, *et al.*, 2001). Hysteresis in radiation and AT for example has been seen multiple times, whilst both variables were found to significantly control net primary production, the largest carbon influx did not occur when they were at their peak, much of the cause of this hysteresis has been attributed to stomatal activity (Huang, *et al.*, 2011, Price and Black, 1990).

3.7. The planetary boundary layer

Figure 13 shows that as expected the PBL deepens during the day, before becoming shallower at night. The daytime depth ranges from 400.7 m to 1074.4 m, whilst most nights see the PBL depth shrink to 100 - 300 m, excluding the 19th – 20th where the PBL became no shallower than

461.7 m. On the 5th and 19th in particular the diurnal PBL cycle has a significantly smaller amplitude than the other days. As aforementioned during daylight hours the anti-correlation between PBL and bioCO₂ is stronger than at night-time, -0.50 and -0.33 respectively. Thus the PBL depth may well exert a stronger control on the bioCO₂ mixing ratio during the day than at night. Whilst diurnal variation of PBL depth does vary quite substantially throughout the study period, there is no homogeneity between large peaks or troughs in the PBL and significantly diurnal bioCO₂ events. Section A for example sees the PBL depth stay very shallow with very little diurnal cycle, and indeed the highest peak in bioCO₂, occurs on the 5th, which could be an effect of the PBL depth. Yet on other days such as the 20th the greatest PBL depth is reached followed by the largest shrinking event of the PBL, occurring simultaneously to section D where there is no noticeable diurnal bioCO₂ cycle. Synonymously, section B sees the largest amplitudes in PBL depth (reaching > 950 m during the day), yet the bioCO₂ diurnal cycles are small.

Further doubt can be placed on the causality between the PBL and the bioCO₂ diurnal variability because previous studies have shown that the CO₂ molar fraction exhibits an asymmetrical shape, with a sharp peak in the CO₂ molar ratio in the morning followed by an elongated trough as shown in Figure 13A (Denning, *et al.*, 1996, Yi, *et al.*, 2000). This occurs even though the flux of CO₂ into and out of the biosphere is roughly symmetrical during the day and night respectively (Baker, *et al.*, 2003). The sharp peak is caused by over-night respiration in the shallow boundary layer followed by a sudden decrease into the elongated trough due to the sudden onset of turbulent mixing and PBL expansion upon sunrise (Yi, *et al.*, 2000). The average shape of both bioCO₂ and obsCO₂ from MHD are demonstrated in Figure 13B, they are far more symmetrical than one would expect, initially suggesting that the PBL does not have a large influence on the diurnal CO₂ or bioCO₂ cycle.

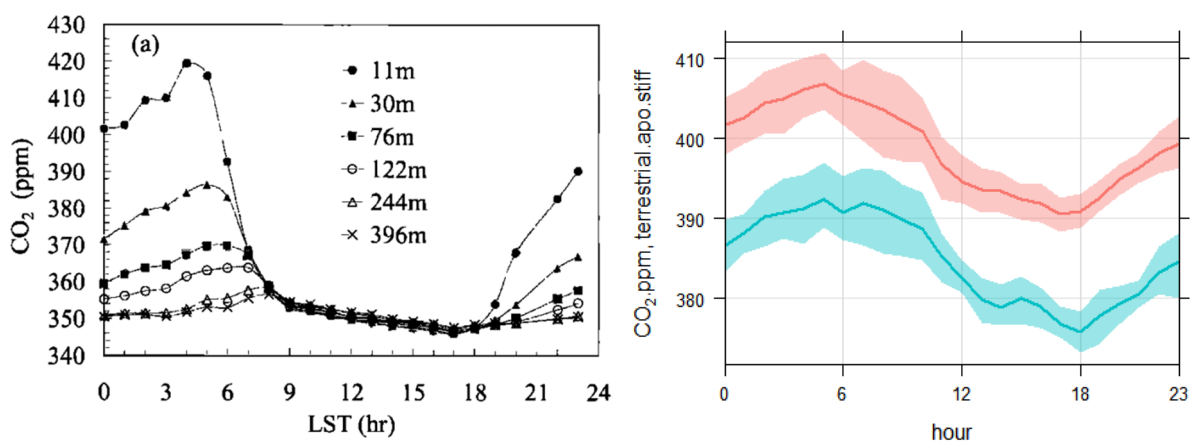


Figure 13 – A) The rectifier effect, it is far more prominent near the surface (Yi, *et al.*, 2000, fig.2) . B) The almost symmetrical bioCO₂ concentration, with the standard deviation represented by the shading.

There are large uncertainties involved when generating data on PBL depth, whilst no uncertainty

was provided with the GDAS model, PBL depth is notoriously hard to determine. Hennemuth & Lammert. (2006) found a deviation of ± 200 m between different methods of determining the PBL. Assuming a similar deviation within which the true PBL depth could lie, dramatic changes could be produced in the correlations, or identifiable links between its depth and PBL seen. Furthermore as the temporal resolution is three hourly it may omit some key features in the PBL activity; especially timing of the peaks and troughs of PBL depth.

3.8. Effect of meteorological variables on the PBL

It is clear that many of the meteorological variables which covary with bioCO₂ do the same with the PBL depth. It is important to keep in mind that these meteorological variables might not just influence the terrestrial biosphere, but may exhibit their effect by influencing the PBL depth (Figure 3); even though it initially appears that the PBL depth is not one of the most significant factors controlling bioCO₂. Radiation is strongly correlated with the PBL, maybe enhancing bioCO₂ significantly through its effect on PBL depth. It is also highly possible that much of the diurnal cycle of RH is a results of the PBL, especially as the strong daylight anti-correlation between the PBL and RH is greater than the RH:bioCO₂ correlation. This may be the result of the expanding and shrinking PBL which is known to decrease and increase RH respectively (Baldocchi, 1997). Without further analysis it is important not to imply causality between any of the variables and bioCO₂, it is indeed possible that the variables covary independently of each other even though there is strong theoretical framework between their relationships.

3.9. The non-correlated meteorological parameters

Whilst AP, CC and precipitation are not directly correlated with bioCO₂ molar fraction they may well still be influential, or at least be indicative of an effect on the biosphere's activity. Direct effects of AP on the terrestrial biosphere cannot be seen during this study as atmospheric pressure does not vary enough (Figure 14). Under artificially controlled conditions it takes increases of 100s of mbar before an affect is seen on the rate of photosynthesis, and as such will never be seen in our atmosphere (Takeishi, *et al.*, 2013). On the other hand, AP can be used as an indicator of changing weather systems, which in turn can be indicative of increased AT, RH or solar radiation, having an effect on synoptic time scales (3 – 7 days) (Baldocchi, *et al.*, 2001). Whilst there are significant changes in the APO over the study period no relationships between AP and bioCO₂ can be attributed.

There were three pulse events of precipitation throughout the study period (Figure 14), and these are not correlated with significant changes in bioCO₂ influx or efflux. It is likely that there is no

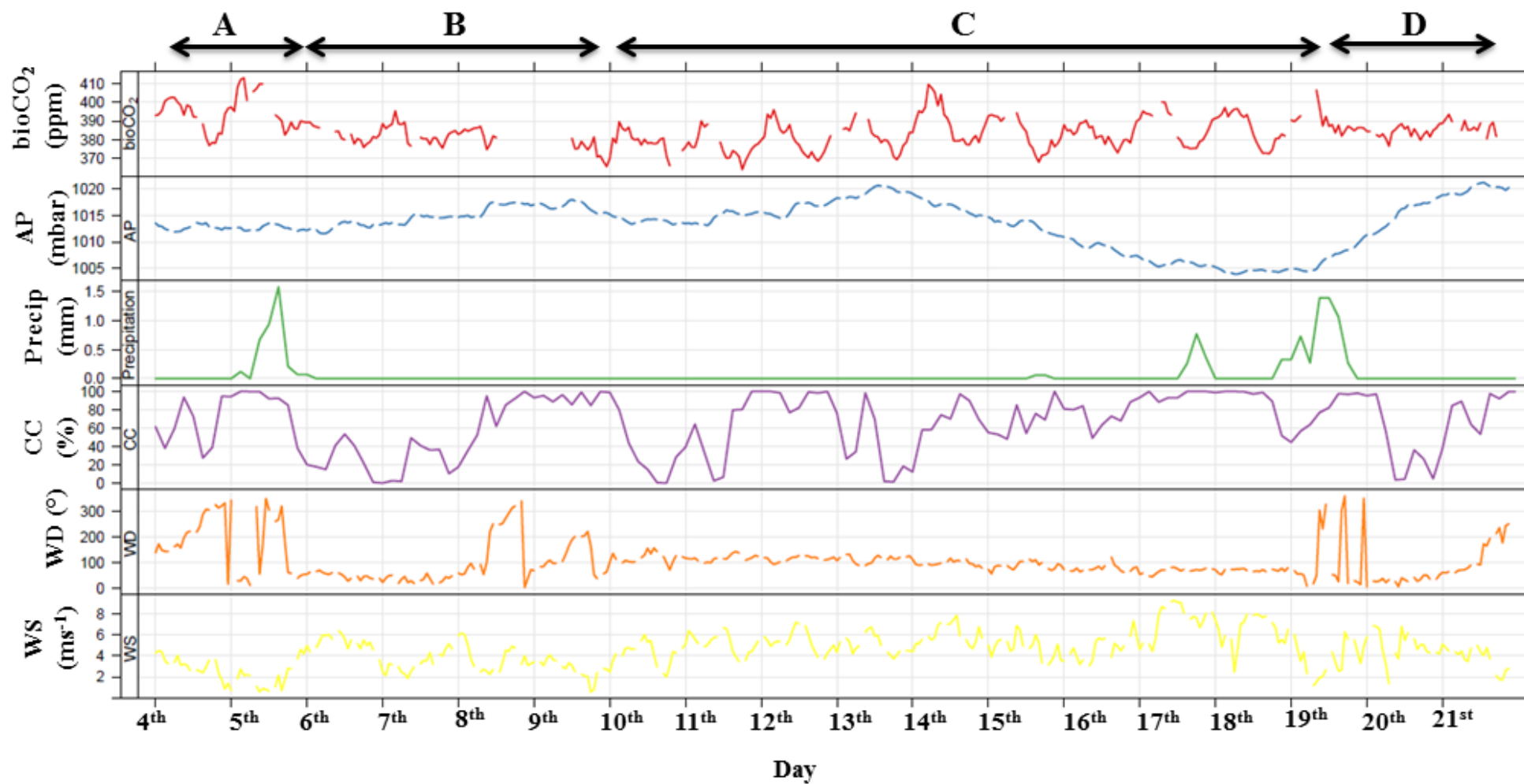


Figure 14 – The meteorological variables that did not correlate with bioCO_2 : AP, precipitation, CC, WD and WS. The pre-defined sections are identified at the top of the graph.

water stress at MHD, in semi-arid ecosystems and even temperate deciduous woodland under water stress precipitation events have shown to increase soil respiration and subsequently carbon efflux to the atmosphere (Lee, *et al.*, 2002). Larger pulses of precipitation (> 5 mm) which provide enough water to penetrate deep into the soil could well act to increase photosynthesis as well as respiration as has been observed in a semi-arid ecosystem and even drought struck peatlands (Huxman, *et al.*, 2004, Kuiper, *et al.*, 2014).

CC appeared to have no significantly determinable effect over bioCO₂ diurnal activity, much of this may be because of the short length of the study period, and furthermore, large changes in CC occurred quite rapidly, making it harder to attribute any relationship between the two. Having said this, clouds can act to increase diffuse radiation, which has been proven to increase photosynthetic light use efficiency in plant canopies and meadow ecosystems (Gu, *et al.*, 2002, Gu, *et al.*, 2003).

3.10 Air mass history – wind speed and direction

The land and sea breeze data, identified in Figure 15, suggests that the net increase and decrease of bioCO₂ in section A has strong local influence. The 7th, 8th, 9th, 10th and 19th all see local influence in the evenings when net drawdown is occurring. As the local influence is not prominent for any significant amount of time no judgement can be made about its effect on the magnitude/timing of the local terrestrial activity. From the short periods of local data that can be seen there is counteracting evidence, for example its presence on the 4th-5th and 19th when the amplitude of the diurnal cycle is large is antagonistic to that on the 7th, 8th, 9th and 10th when the diurnal cycles are small. The vast majority of section C, where the diurnal variability is generally largest and at its most defined is unaffected by breezes and as such unlikely to be affected by local influence.

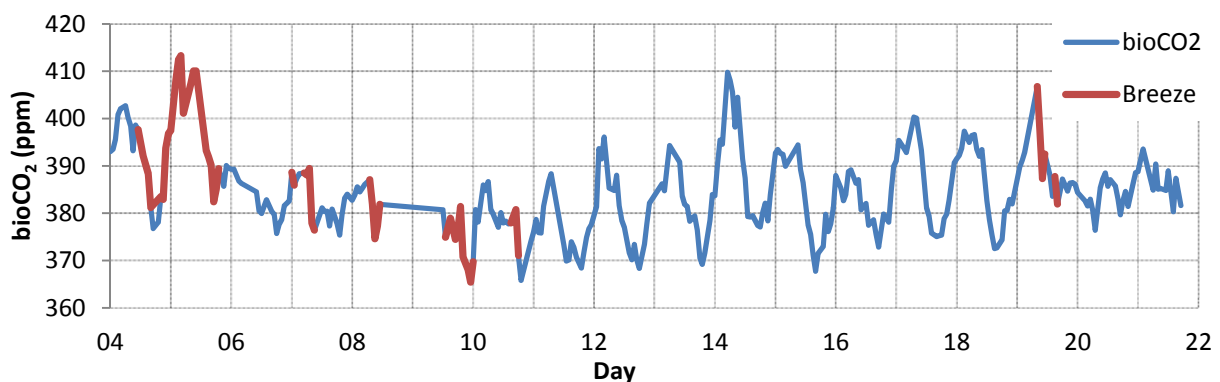


Figure 15 - The data with a wind speed $< 3 \text{ ms}^{-1}$ is plotted in red, over the bioCO₂ data for the study period. Large marker points have been implemented into the breeze data to ensure single points of slow wind speed are visible.

Figure 14 shows distinct differences between the wind regimes in section A, B, C and D. Section A sees strong fluctuations in the WD which appear to be associated with local origin. Over much of section B where the diurnal cycle is small vast amounts of the wind approaches from 40 –

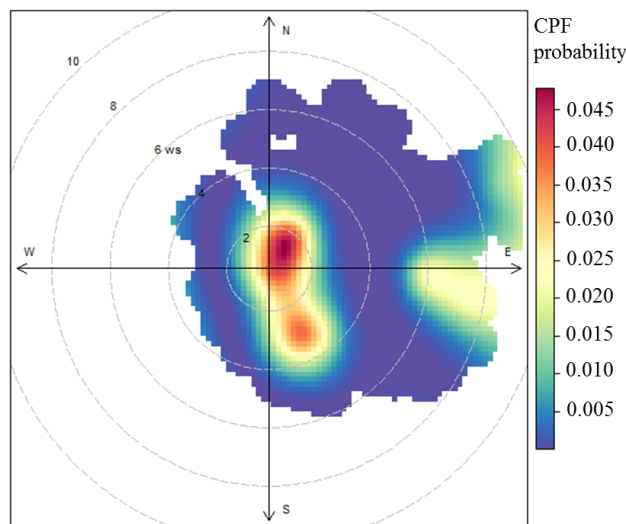


Figure 17 – CPF plot for September, covering bioCO₂ values in the 95th-100th percentile (400 – 413 ppm)

60 °, suggesting North-East source location, perhaps N Ireland, Scotland, the North Sea and maybe even Northern Europe. Section C sees the most uniformly high WS and distinct easterly WD which stays within the bounds of 100 ° to 60 °, slowly rotating from the former to the latter from the 10th to the 19th, this signal could be attributed to air masses from Ireland, the UK and even mainland Europe. Interestingly, section D where the diurnal cycle is at its smallest and noisiest shows wind direction which shifts from Easterlies to Westerlies, where the air starts to approach

MHD off of the Atlantic. Throughout the study period there are several spikes in the WD where the wind goes above 300 °, before transferring back to easterlies, it appears that these may coincide with the breeze data and strong influence from the local terrestrial biosphere.

The CPF plot in Figure 17 further demonstrates what was seen, the largest bioCO₂ values being most likely to arrive from Easterly winds with a slow WS, representative of local activity, most likely influenced by section A. It is next most likely that high bioCO₂ molar fractioned air arrives from an easterly direction with WS > 6 ms⁻¹, which is influenced by section C.

3.11. Air mass history – back trajectories

It is clear from the Figure 18 that the transport of air masses has a defined effect on the bioCO₂ signal at MHD. The bioCO₂ diurnal cycle has larger amplitude when the air masses have originated from Northern Europe or passed through the British Isles than those originating directly from the Atlantic. Sections B and D clearly show that the air arrives explicitly from the Atlantic, along trajectories in clusters C4 and C5. Section C on the other hand which has the most defined bioCO₂ cycle sees air arriving from South Scandinavia (C2 and C6) and the North Atlantic, passing through the British Isles (C1). Section A does see air masses arriving from the British Isles, around Ireland and Northern Europe (C1, C3 and C2). The exact timings of the back trajectories can be found in the Appendix, Figure 3 & Table 1.

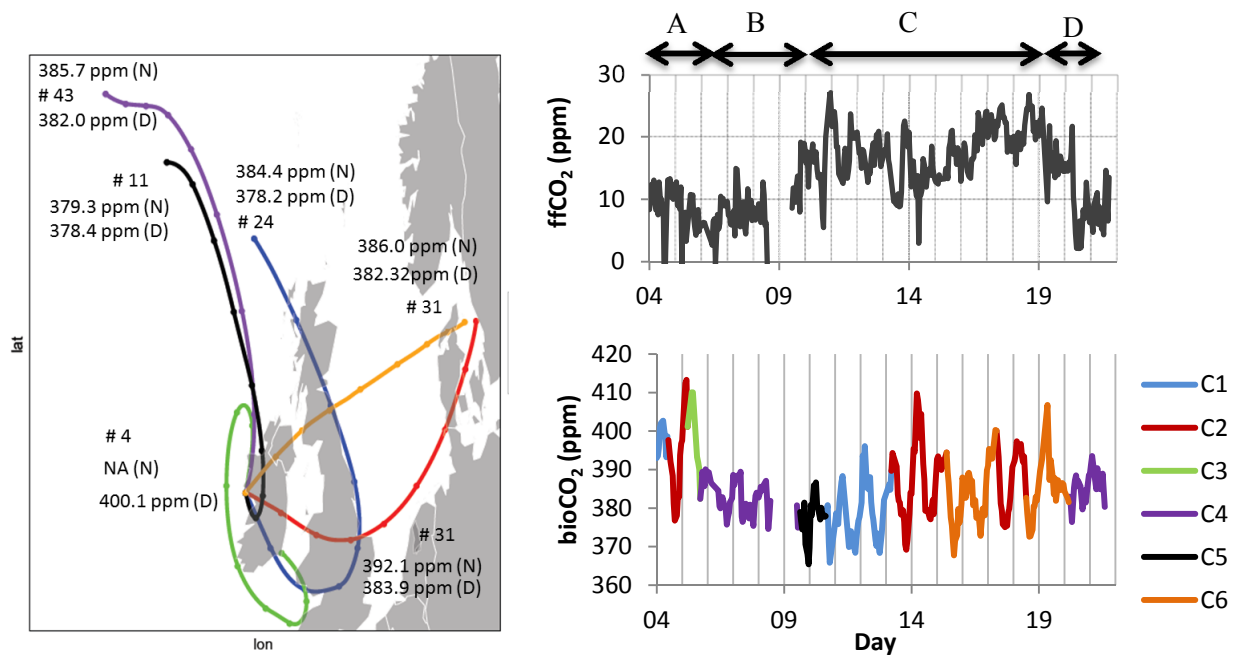


Figure 18 – The back trajectories for the entire September study period, separated into six clusters. The number of back trajectories in each cluster is represented by #, whilst the average daylight (D) and night-time (N) CO₂ molar fractions are shown for each cluster. There were no night-time back trajectories for C3.

Section A could see its large respiration signal influenced by the Northern European origin of air mass history and large photosynthesis signal influenced by much of the British Isles (C3). It is however far more likely that this is due to local influence, not just because local breeze was occurring when these measurements were made, but because the ffCO₂ data for this time period is low. As simulated using REMO and EDGAR emissions data (shown in Figure 19), the emissions of Ireland and the region local to MHD are weak, local origin therefore should and does occur simultaneously to the lower ffCO₂ values. In section C when the signal is derived from Northern Europe there is a higher average ffCO₂ signal. Figure 18 again shows just how strong the ffCO₂ signal is over the Jutland region and British Isles both of which are likely to have influenced the air mass composition measured at MHD. Sections B and D, similarly to section A also see smaller average ffCO₂ molar fractions, the vast majority of which will be associated with local emissions.

It is clear from this data that the smallest, least defined bioCO₂ cycles have an oceanic origin. This is part due to ocean-carbonate chemistry, dissolved CO₂ equilibrates with various carbon species, carbonate and bicarbonate ions in particular, acting to suppress the marine photosynthesis and respiration signal of dissolved CO₂ 15 fold (Keeling and Shertz, 1992). Furthermore it takes ~1 year for the partial pressure of CO₂ to equilibrate between the ocean and the atmosphere, both of these effects act to completely dissipate any diurnal terrestrial signal from forming above the ocean (Figure 19B) (Keeling and Shertz, 1992). The small diurnal bioCO₂ cycles seen in section B and D must be due solely to the activity of the local biosphere,

and signals which have originated from Ireland on its trajectory after passing over land, as aforementioned this is supported by the relatively low (but existent) ffCO₂ molar fraction.

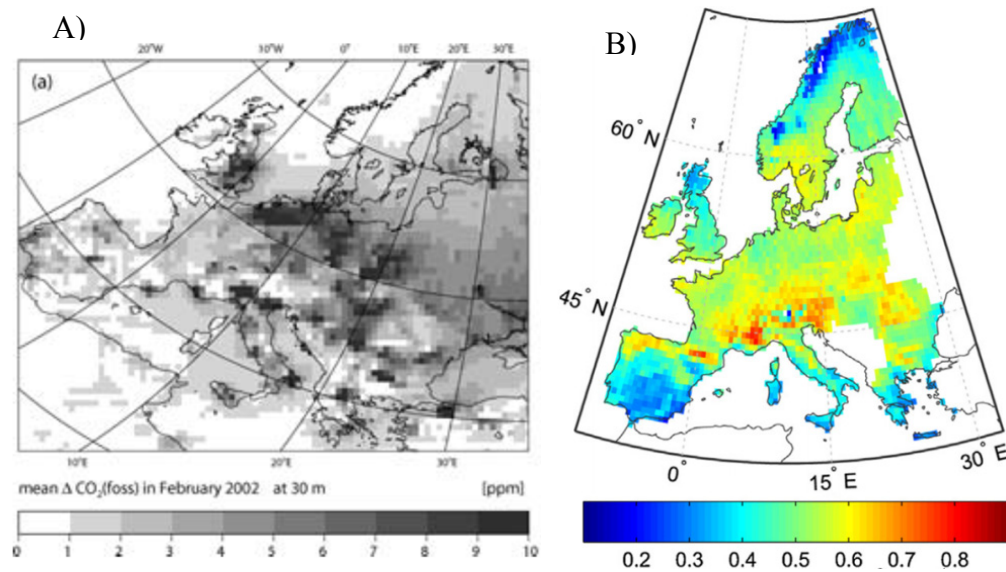


Figure 19 –A) The ffCO₂ signal over Europe as simulated by REMO, using EDGAR 30 m above ground (Levin and Karstens, 2007) fig.2. B) NPP of Europe as simulated by the LPJ-GUESS dynamic vegetation model (KgC m⁻² year⁻¹) (Sallaba, *et al.*, 2015) fig.5.

Air masses from C1, C2 and C6 presumably bring the land biotic signal with them from mainland Europe as well as from the British Isles. As shown by the map of net primary productivity (NPP) for Europe (Figure 19) it is clear that south Scandinavia has a highly active terrestrial biosphere, and so does South Ireland, explaining why the bioCO₂ diurnal signal is strong for both for air masses which have originated from or passed these locations. If air masses were to originate from a more southerly location of the European continent such as North Italy or Switzerland location the bioCO₂ activity may be even larger.

The CWT plot in Figure 21 can be used to interpret possible bioCO₂ source locations which have contributed to the bioCO₂ signal seen at MHD (Cheng, *et al.*, 2013). Areas shaded in red represent the largest contribution, whilst those in blue represent the smallest contribution to the total signal. It is clear that the bioCO₂ source locations are predominantly from Ireland, with Wales and South England supplying contributing the next most to the bioCO₂ signal. This suggests that whilst the air masses may have arrived at MHD from as far afield as Southern Scandinavia, the vast majority of the contributions to the bioCO₂ signal seen at MHD are from the more local sources. Trajectories in C2 and C6 may well see large amounts of the bioCO₂ signal originate from Ireland, even though the actual trajectory originates from the highly polluted and terrestrially active South Scandinavia. This can be further backed up the trajectories in C1 which originate over the Atlantic, but are associated with a large bioCO₂ diurnal cycle

which must be incident from the British Isles.

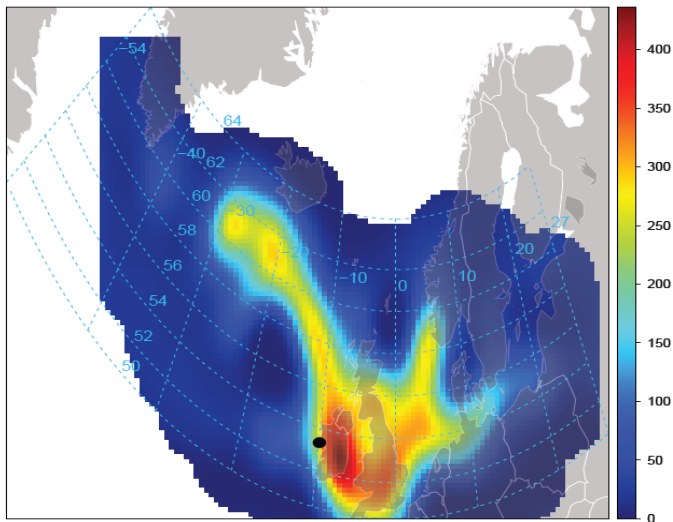


Figure 21– Concentration field plot of bioCO₂ over the September study period.

3.12. Caveats in the ffCO₂ methodology and areas for future research

It is clear from the uncertainty analysis that the baseline and emissions ratios provide significant uncertainties, vast improvements of which will be required in the future. A universal gage of the correct span for the REBS baseline should be identified to generate the most accurate flexibility of the baseline. Alternatively, at MHD the baselines could have been produced by using back trajectories or the backwards function of a dispersion model, guaranteeing that the synoptic scale transport of air is attributed to Atlantic origin. Whilst dispersion models are similar to back trajectories, they allow singular particles to move independently from each other whilst taking into account turbulence and vertical mixing (Fleming, *et al.*, 2012). Both of these methods have been used at MHD, and show good results (Derwent, *et al.*, 1998, Ryall, *et al.*, 2001). Their complexity however means they could not be used under the given time constraints, furthermore dispersion models require the use of a super computer.

Similar studies in the future will require the COFFEE and EDGAR inventories to be updated so that they temporally match the data sets. It may also be desirable to incorporate emissions ratios from other inventories such as IER for R_{CO}, to my knowledge COFFEE is the only inventory which can provide R_{AP0}. Whilst these factors are reliant on the inventory production process, the use of backward functions of dispersion models as opposed to back trajectories could be used to extract the emissions ratios more accurately. Using a dispersion model the weighting of different locations contributing to the air reaching the receptor site could be determined, and as such one would expect the ratios to be extracted more accurately than by the trajectory-based method used here.

3.13. Further areas of research

As factors such as hysteresis and time lags act to convolute the effect of environmental variables on the local biosphere it is clear that a far longer study period would be required to identify causality between meteorological data and the local biosphere. Baldocchi *et al.* (2001) used spectral analysis to mathematically identify complex temporal patterns in the exchange of trace gases between a deciduous woodland and the atmosphere. They identified relations between different variables such as temperature and sunlight which were not readily detectable from the time series data. This mathematical technique may well be an invaluable if a long enough data set can be produced at MHD. It would also be highly desirable to obtain data on a wider range of variables such as SM, ST and soil nutrient content, all of which have been shown to play a significant role in determining the activity of the terrestrial biosphere (Urbanski, *et al.*, 2007).

As there is also convolution between the effects of vertical mixing and the terrestrial biosphere activity, it would be very informative to calculate biosphere fluxes (*bioCO₂), so that meteorological variables can be associated explicitly to the terrestrial biosphere activity. The eddy covariance technique is widely used to decipher carbon fluxes at an ecosystem level, but would require a tall-tower measurement centre (Guan, *et al.*, 2005). Alternatively radon (²²²Rn) emissions to atmospheric molar fraction can be used to quantify local emission fluxes for short term events (Van der Laan, *et al.*, 2010). This would be based on ²²²Rn fluxes being well constrained at MHD, and the implementation of new measurement equipment (Levin, *et al.*, 1989).

On the other hand, inversion modelling can be used identically to determine *bioCO₂ on a global to regional scale; ²²²Rn data is then used to correct the model for errors in its transport calculations, specifically the vertical mixing (Vogel, *et al.*, 2013). Regional CO₂ fluxes to/from the terrestrial biosphere could also be estimated by process models such as LandscapeDNDC to obtain a better understanding of the CO₂ source and sink regions surrounding the study site (Molina-Herrera, *et al.*, 2015).

Combining efforts from the above studies could be used to better decipher the variables which control the bioCO₂ diurnal variation. On top of this, the use of a dispersion model when determining air mass history would provide a superior image of the locations which have contributed air to the receptor site compared to back trajectories (Ryall, *et al.*, 2001); potentially allowing further attributions to be made.

4. Conclusion

Estimates of the atmospheric molar fraction of ffCO₂ at MHD from February 2014 – July 2015 were more precise when quantified using the novel ffCO₂ tracer APO than CO, even though the O₂ measurements were not to the desirable precision. When quantifying ffCO₂ using CO or APO there are three sources of uncertainty: the measurements, baseline and the emissions ratio. The uncertainty for each was higher for the CO methodology than the APO methodology. It is however, the emissions ratio which provides the main uncertainty for the CO method, reaching an uncertainty of 303.6 %, far larger than the largest uncertainty of 85.2 % obtained for the APO method. Whilst it seems highly likely that much of this uncertainty is due to erroneous figures in the EDGAR inventory, the range of values that R_{APO} can hold is far narrower than R_{CO}. Ranges 0.22 per mer ppm⁻¹ and 118.53 ppb ppm⁻¹ were seen in this study, and as such inaccuracies in R_{CO} can be transcribed into far larger inaccuracies of the ffCO₂ values.

Whilst the accuracy of using APO as a ffCO₂ tracer has not been quantified here, it is affected by far fewer biases, with variation in its molar fraction on short time scales at least being biased only by liquid and gaseous biofuels. Furthermore, initial ffCO₂ results suggested that the values generated by APO were more realistic of the rural location studied. These initial tests certainly suggest that APO could be far more accurate and precise as a tracer than CO, a positive outcome which could eventually see the inclusion of top-down emission estimates in emissions reporting. It is clear from this study that there are still some significant uncertainties associated with using APO. Maybe the largest drawback associated with APO its use is the bias due to liquid and gaseous biofuel emissions, which may become an ever increasing energy source in the future. In regions without nuclear influence, ¹⁴C should still be considered the gold standard for ffCO₂ quantification, and could be used to calibrate the ffCO₂(APO) to produced datasets with a high temporal resolution.

At MHD CO₂ fluctuated diurnally in September 2014, this saw the CO₂ molar fraction increase overnight, reaching a peak in the early morning before increasing during the day; a pattern which can be attributed almost solely to terrestrial biosphere photosynthesis and respiration. The recorded bioCO₂ signal is however, influenced most clearly by the synoptic scale transport of air masses reaching MHD. Air arriving directly from the Atlantic is associated with the smaller diurnal bioCO₂ cycle, whilst air incident from mainland Europe and those which have originated from the Atlantic but passed through the British Isles can be associated with the diurnal cycles of bioCO₂ with the largest amplitude. Whilst it is entirely possible that some of the signal has arrived from the source regions of the trajectories or further, it is more likely that the bioCO₂

signal seen has been picked up by the air mass relatively close to MHD, as it passed through the British Isles, in particular Ireland, Wales and South England.

Local meteorological conditions, in particular RH, radiation and AT were moderately to strongly correlated/anti-correlated with bioCO₂ and preceded its activity, suggesting a cause and effect relationship. Quantification of the effect of each variable was not possible as no variable appeared to have any significant influence on the bioCO₂ molar ratios. No causality could therefore be established, probably due to hysteresis, chaos, time lags and filtering, all of which have been identified in previous ecosystem studies. None of the meteorological variables appeared to have a strong individual influence on bioCO₂, indicating that they are not a limiting factor for respiration or photosynthesis. It seems most likely that these variables act to control the terrestrial biosphere in a summative manner, each one contributing to the rate of photosynthesis or respiration. Alternatively, as causality could not be established it is possible that these parameter are just covarying in an independent manner.

Whilst it is well known that trace gas molar fractions are diluted when the PBL is deep, and see increased molar fractions when it is shallow, only moderate correlations were obtained here. In fact the shape of the diurnal CO₂ molar fraction was not indicative of any PBL influence at all. Due to the simplistic nature of this study, the effect of PBL depth could not be quantified; this is certainly an area for further research.

In this study inferences have been made about causing factors behind the variation bioCO₂ signal at MHD on a diurnal time scale, but these are by no means conclusive. It is clear that in future studies the local influence from the terrestrial biosphere should be separated out from atmospheric transport to reduce the amount of convoluting factors. Much more research will be required in the future to decipher the causes of variability in the bioCO₂ cycle taking into account a wider range of environmental variables over a far greater time period. Ecosystem scale studies using the eddy covariance technique might well be necessary to better understand the local behaviour of the biosphere, whilst transport models with ²²²Rn correction and process models could be used to better quantify the sources and sinks attributing to the MHD bioCO₂ signal. In particular, model correction using ²²²Rn would be highly informative when trying to decipher the effect of PBL. It is vital that studies such as this are carried out in the future, as our understanding of the terrestrial biosphere is rather limited. Understanding of its ever-changing activity as a sink or source is vital to understand the climate sensitivity so that future climate change projected, and and act accordingly to mitigate it.

References

1. I. Alados and L. Alados-Arboledas, 1999, Direct and diffuse photosynthetically active radiation: measurements and modelling, *Agricultural and Forest Meteorology*, **93**, 27-38.
2. R. J. Andres, *et al.*, 2011, Monthly, global emissions of carbon dioxide from fossil fuel consumption, *Tellus Series B-Chemical and Physical Meteorology*, **63**, 309-327.
3. I. T. Baker, *et al.*, 2003, Simulated and observed fluxes of sensible and latent heat and CO₂ at the WLEF-TV tower using SiB2.5., *Global Change Biology*, **9**, 1262 - 1277.
4. P. S. Bakwin, P. P. Tans, D. F. Hurst and C. L. Zhao, 1998, Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program, *Tellus Series B-Chemical and Physical Meteorology*, **50**, 401-415.
5. D. Baldocchi, 1997, Measuring and modelling carbon dioxide and water vapour exchange over a temperate broad-leaved forest during the 1995 summer drought, *Plant Cell and Environment*, **20**, 1108-1122.
6. D. Baldocchi, E. Falge and K. Wilson, 2001, A spectral analysis of biosphere-atmosphere trace gas flux densities and meteorological variables across hour to multi-year time scales, *Agricultural and Forest Meteorology*, **107**, 1-27.
7. T. W. Blaine, R. F. Keeling and W. J. Paplawsky, 2005, An improved inlet for precisely measuring the atmospheric Ar/N₂ ratio, *Atmos. Chem. Phys.*, **6**, 1181-1184.
8. K. Bouchlaghem, F. Ben Mansour and S. Elouragini, 2007, Impact of a sea breeze event on air pollution at the Eastern Tunisian Coast, *Atmospheric Research*, **86**, 162-172.
9. P. Bousquet, *et al.*, 1996, Atmospheric CO₂ concentration variations recorded at Mace Head, Ireland, from 1992 to 1994, *Physics and Chemistry of the Earth*, **21**, 477 - 481.
10. J. S. Boyer, 2015, Turgor and the transport of CO₂ and water across the cuticle (epidermis) of leaves, *Journal of Experimental Botany*, **66**, 2625-2633.
11. D. Bozhinova, *et al.*, 2014, Simulating the integrated summertime $\Delta^{14}\text{CO}_2$ signature from anthropogenic emissions over Western Europe, *Atmospheric Chemistry and Physics*, **14**, 7273-7290.
12. J. G. Canadell, *et al.*, 2007, Contributions to accelerating atmospheric CO₂ growth from economic activity, carbon intensity, and efficiency of natural sinks, *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 18866-18870.
13. J. N. Cape, J. Methven and L. E. Hudson, 2000, The use of trajectory cluster analysis to interpret trace gas measurements at Mace Head, Ireland, *Atmospheric Environment*, **34**, 3651 - 3663.
14. D. C. Carslaw and K. Ropkins, 2012, openair - An R package for air quality data analysis, *Environmental Modelling & Software*, **27-28**, 52-61.

15. I. Cheng, *et al.*, 2013, Concentration-weighted trajectory approach to identifying potential sources of speciated atmospheric mercury at an urban coastal site in Nova Scotia, Canada, *Atmospheric Chemistry and Physics*, **13**, 6031-6048.
16. A. S. Denning, I. Y. Fung and D. A. Randall, 1995, Latitudinal gradient of atmospheric CO₂ due to seasonal exchange with land biota, *Nature*, **376**, 240-243.
17. A. S. Denning, D. Randall, G. J. Collatz and P. J. Sellers, 1996, Simulations of terrestrial carbon metabolism and atmospheric CO₂ in a general circulation model, *Tellus*, **48B**, 543 - 567.
18. R. G. Derwent, *et al.*, 2002, Continuous observations of carbon dioxide at Mace Head, Ireland from 1995 to 1999 and its net European ecosystem exchange, *Atmospheric Environment*, **36**, 2799-2807.
19. R. G. Derwent, P. G. Simmonds, S. Seuring and C. Dimmer, 1998, Observation and interpretation of the seasonal cycles in the surface concentrations of ozone and carbon monoxide at Mace Head, Ireland from 1990 to 1994, *Atmospheric Environment*, **32**, 145-157.
20. R. R. Draxler and G. D. Hess, 1998, An overview of the HYSPLIT_4 modelling system for trajectories, dispersion, and deposition, *Aust. Met. Mag.*, **47**, 295-308.
21. M. Ek and L. Mahrt, 1994, Daytime evolution of relative humidity at the boundary-layer top, *Monthly Weather Review*, **122**, 2709-2721.
22. Z. L. Fleming, P. S. Monks and A. J. Manning, 2012, Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition, *Atmospheric Research*, **104**, 1-39.
23. P. Friedlingstein, *et al.*, 2014, Persistent growth of CO₂ emissions and implications for reaching climate targets, *Nature Geoscience*, **7**, 709-715.
24. P. Friedlingstein, *et al.*, 2006, Climate-carbon cycle feedback analysis: Results from the C⁴MIP model intercomparison, *Journal of Climate*, **19**, 3337-3353.
25. U. Gamnitzer, *et al.*, 2006, Carbon monoxide: A quantitative tracer for fossil fuel CO₂?, *Journal of Geophysical Research*, **11**,
26. C. W. J. Granger, 1988, Some recent developments in a concept of causality, *Journal of Econometrics*, **39**, 199-211.
27. C. Granier, G. Petron, J. F. Muller and G. Brasseur, 2000, The impact of natural and anthropogenic hydrocarbons on the tropospheric budget of carbon monoxide, *Atmospheric Environment*, **34**, 5255-5270.
28. H. D. Graven and N. Gruber, 2011, Continental-scale enrichment of atmospheric ¹⁴CO₂ from the nuclear power industry: potential impact on the estimation of fossil fuel-derived CO₂, *Atmospheric Chemistry and Physics*, **11**, 12339-12349.
29. L. H. Gu, *et al.*, 2002, Advantages of diffuse radiation for terrestrial ecosystem productivity, *Journal of Geophysical Research-Atmospheres*, **107**,

30. S. Gu, *et al.*, 2003, Short-term variation of CO₂ flux in relation to environmental controls in an alpine meadow on the Qinghai-Tibetan Plateau, *Journal of Geophysical Research-Atmospheres*, **108**,
31. D. X. Guan, *et al.*, 2005, Meteorological control on CO₂ flux above broad-leaved Korean pine mixed forest in Changbai Mountains, *Science in China Series D-Earth Sciences*, **48**, 116-122.
32. R. C. Hamme and R. F. Keeling, 2008, Ocean ventilation as a driver of interannual variability in atmospheric potential oxygen, *Tellus Series B-Chemical and Physical Meteorology*, **60**, 706-717.
33. B. Hennemuth and A. Lammert, 2006, Determination of the atmospheric boundary layer height from radiosonde and lidar backscatter, *Boundary-Layer Meteorology*, **120**, 181-200.
34. I. Y. Hernandez-Paniagua, *et al.*, 2015, Diurnal, seasonal, and annual trends in atmospheric CO₂ at southwest London during 2000-2012: Wind sector analysis and comparison with Mace Head, Ireland, *Atmospheric Environment*, **105**, 138-147.
35. H. Huang, *et al.*, 2011, Seasonal variation and meteorological control of CO₂ flux in a hilly plantation in the mountain areas of North China, *Acta Meteorologica Sinica*, **25**, 238-248.
36. T. E. Huxman, *et al.*, 2004, Precipitation pulses and carbon fluxes in semiarid and arid ecosystems, *Oecologia*, **141**, 254-268.
37. IPCC, 2013, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
38. IPCC, 2007, *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [S. Solomon, D. Qin, M. Manning, B. Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA
39. B. H. Jia, *et al.*, 2015, Diurnal and Seasonal Variations of CO₂ Fluxes and Their Climate Controlling Factors for a Subtropical Forest in Ningxiang, *Advances in Atmospheric Sciences*, **32**, 553-564.
40. E. Kalnay, *et al.*, 1996, The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, **77**, 437-471.
41. C. D. Keeling, J. F. S. Chin and T. P. Whorf, 1996, Increased activity of northern vegetation inferred from atmospheric CO₂ measurements, *Nature*, **382**, 146-149.
42. C. D. Keeling, T. P. Whorf, M. Wahlen and J. van der Plicht, 1995, Interannual extremes in the rate of rise of atmospheric carbon dioxide since 1980, *Nature*, **375**, 666-670.
43. R. F. Keeling, 1988, Development of an interferometric oxygen analyzer for precise measurement of the atmospheric O₂ mole fraction, Ph.D thesis, The Division of Applied Sciences, Harvard University, Cambridge, Massachusetts.

44. R. F. Keeling and S. R. Shertz, 1992, Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle, *Nature*, **358**, 723-727, doi:10.1038/358723a0.
45. W. Knorr, 2009, Is the airborne fraction of anthropogenic CO₂ emissions increasing?, *Geophysical Research Letters*, **36**,
46. J. J. Kuiper, W. M. Mooij, L. Bragazza and B. J. M. Robroek, 2014, Plant functional types define magnitude of drought response in peatland CO₂ exchange, *Ecology*, **95**, 123-131.
47. O. L. Lange, R. Lösch, E. D. Schulze and L. Kappen, 1971, Responses of stomate to changes in humidity, *Planta*, **100**, 76 - 86.
48. M. Lee, *et al.*, 2002, Effects of rainfall events on soil CO₂ in a cool temperate deciduous broad-leaved forest, *Ecological Research*, **17**, 401 - 409.
49. I. Levin, S. Hammer, B. Kromer and F. Meinhardt, 2008, Radiocarbon observations in atmospheric CO₂: Determining fossil fuel CO₂ over Europe using Jungfraujoch observations as background, *Science of the Total Environment*, **391**, 211-216.
50. I. Levin and U. Karstens, 2007, Inferring high-resolution fossil fuel CO₂ records at continental sites from combined ¹⁴CO₂ and CO observations, *Tellus B*, **59**, 245-250.
51. I. Levin, B. Kromer, M. Schmidt and H. Sartorius, 2003, A novel approach for independent budgeting of fossil fuel CO₂ over Europe by ¹⁴CO₂ observations, *Geophysical Research Letters*, **30**,
52. I. Levin, J. Schuchard, B. Kromer and K. O. Munnich, 1989, The continental European Suess effect, *Radiocarbon*, **31**, 431-440.
53. M. Lopez, *et al.*, 2013, CO, NO_x and ¹³CO₂ as tracers for fossil fuel CO₂: results from a pilot study in Paris during winter 2010, *Atmospheric Chemistry and Physics*, **13**, 7343-7358.
54. A. J. Manning, 2011, The challenge of estimating regional trace gas emissions from atmospheric observations, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **369**, 1943-1954.
55. P. C. Mielnick and W. A. Dugas, 2000, Soil CO₂ flux in a tallgrass prairie, *Soil Biology & Biochemistry*, **32**, 221-228.
56. Y. Miyaoka, *et al.*, 2007, Diurnal and seasonal variations in atmospheric CO₂ in Sapporo, Japan: Anthropogenic sources and biogenic sinks, *Geochemical Journal*, **41**, 429-436.
57. S. Molina-Herrera, *et al.*, 2015, Simulation of CO₂ Fluxes in European Forest Ecosystems with the Coupled Soil-Vegetation Process Model "LandscapeDNDC", *Forests*, **6**, 1779-1809.
58. J. L. Moody, J. A. Galusky and J. N. Galloway, 1989, The use of atmospheric transport pattern recognition techniques in understanding variation in precipitation chemistry, *Atmospheric Deposition*, **179**, 119 - 125.
59. E. Nisbet and R. Weiss, 2010, Top-down versus bottom-up, *Science*, **328**, 1241-1243.
60. S. O'Doherty, 2015, CO. Bristol, Unpublished data.

61. P. A. Pickers, 2016, New applications of continuous atmospheric O₂ measurements: meridional transects across the Atlantic Ocean, and improved quantification of fossil fuel-derived CO₂, Ph.D thesis, in prep, University of East Anglia.
62. D. T. Price and T. A. Black, 1990, Effects of short-term variation in weather on diurnal canopy CO₂ flux and evapotranspiration of a juvenile Douglas-fir stand, *Agricultural and Forest Meteorology*, **50**, 139-158.
63. M. Ramonet, 2015, P. Laboratoire des Sciences du Climat et de l'Environnement (LSCE), France, unpublished data.
64. M. Ramonet and P. Monfray, 1996, CO₂ baseline concept in 3-D atmospheric transport models, *Tellus*, **48B**, 502 - 520.
65. R. M. Rotty, 1987, Estimates of seasonal variation in fossil fuel CO₂ emissions, *Tellus*, **39B**, 184 - 202.
66. A. F. Ruckstuhl, *et al.*, 2012, Robust extraction of baseline signal of atmospheric trace species using local regression, *Atmospheric Measurement Techniques*, **5**, 2613-2624.
67. A. F. Ruckstuhl and R. Locher, 2012, IDPmisc: Utilities of institute of data analyses and process design, R package version 1.1.17.
68. D. B. Ryall, *et al.*, 2001, Estimating source regions of European emissions of trace gases from observations at Mace Head, *Atmospheric Environment*, **35**, 2507-2523.
69. G. Salisbury, *et al.*, 2002, A seasonal comparison of the ozone photochemistry in clean and polluted air masses at Mace Head, Ireland, *Journal of Atmospheric Chemistry*, **41**, 163-187.
70. F. Sallaba, D. Lehsten, J. Seaquist and M. T. Sykes, 2015, A rapid NPP meta-model for current and future climate and CO₂ scenarios in Europe, *Ecological Modelling*, **302**, 29-41.
71. D. S. Schimel, *et al.*, 2001, Recent patterns and mechanisms of carbon exchange by terrestrial ecosystems, *Nature*, **414**, 169-172.
72. H. Schoh, M. Bruns, K. O. Munnich and M. Munnich, 1980, A multi-counter system for high precision carbon-14 measurements, *Radiocarbon*, **22**, 442 - 447.
73. P. G. Simmonds, S. Seuring, G. Nickless and R. G. Derwent, 1997, Segregation and interpretation of ozone and carbon monoxide measurements by air mass origin at the TOR Station Mace Head, Ireland from 1987 to 1995, *Journal of Atmospheric Chemistry*, **28**, 45-59.
74. R. E. H. Sims, *et al.*, 2006, Energy crops: current status and future prospects, *Global Change Biology*, **12**, 2054-2076.
75. J. Steinbach, *et al.*, 2011, The CO₂ release and Oxygen uptake from Fossil Fuel Emission Estimate (COFFEE) dataset: effects from varying oxidative ratios, *Atmos. Chem. Phys.*, **11**, 6855-6870.
76. B. B. Stephens, *et al.*, 1998, Testing global ocean carbon cycle models using measurements of atmospheric O₂ and CO₂ concentration, *Global Biogeochemical Cycles*, **12**, 213-230.

77. M. Stuiver and H. A. Polach, 1977, Reporting of ^{14}C data, *Radiocarbon*, **19**, 355 - 363.
78. H. E. Suess, 1955, Radiocarbon concentration in modern wood, *Science*, **122**, 415-417.
79. H. Takeishi, *et al.*, 2013, Effects of elevated pressure on rate of photosynthesis during plant growth, *Journal of Biotechnology*, **168**, 135-141.
80. R. D. C. Team, 2008, R: A language and environment for statistical computing,
81. J. C. Turnbull, *et al.*, 2006, Comparison of $^{14}\text{CO}_2$, CO, and SF_6 as tracers for recently added fossil fuel CO_2 in the atmosphere and implications for biological CO_2 exchange, *Geophysical Research Letters*, **33**, 5.
82. S. Urbanski, *et al.*, 2007, Factors controlling CO_2 exchange on timescales from hourly to decadal at Harvard Forest, *Journal of Geophysical Research*, **112**,
83. S. Van der Laan, *et al.*, 2010, Observation-based estimates of fossil fuel-derived CO_2 emissions in the Netherlands using $\Delta^{14}\text{C}$, CO and ^{222}Rn , *Tellus Series B-Chemical and Physical Meteorology*, **62**, 389-402.
84. S. N. Vardag, C. Gerbig, G. Janssens-Maenhout and I. Levin, 2015, Estimation of continuous anthropogenic CO_2 : model-based evaluation of CO_2 , CO, $\delta^{13}\text{C}(\text{CO}_2)$ and $\Delta^{14}\text{C}(\text{CO}_2)$ tracer methods, *Atmospheric Chemistry and Physics*, **15**, 12705-12729.
85. F. R. Vogel, *et al.*, 2010, Implication of weekly and diurnal ^{14}C calibration on hourly estimates of CO-based fossil fuel CO_2 at a moderately polluted site in southwestern Germany, *Tellus Series B-Chemical and Physical Meteorology*, **62**, 512-520.
86. F. R. Vogel, *et al.*, 2013, Can we evaluate a fine-grained emission model using high-resolution atmospheric transport modelling and regional fossil fuel CO_2 observations?, *Tellus Series B-Chemical and Physical Meteorology*, **65**, 16.
87. G. L. Vourlitis, *et al.*, 2004, Effects of meteorological variations on the CO_2 exchange of a Brazilian transitional tropical forest, *Ecological Applications*, **14**, S89-S100.
88. J. M. Waddington, P. A. Rotenberg and F. J. Warren, 2001, Peat CO_2 production in a natural and cutover peatland: Implications for restoration, *Biogeochemistry*, **54**, 115-130.
89. P. Wilson, 2012, Insight into the carbon cycle from continuous measurements of Oxygen and Carbon Dioxide at Weybourne atmospheric observatory, UK, Ph.D, School of Environmental Sciences, University of East Anglia.
90. S. H. Wu, P. E. Jansson and P. Kolari, 2012, The role of air and soil temperature in the seasonality of photosynthesis and transpiration in a boreal Scots pine ecosystem, *Agricultural and Forest Meteorology*, **156**, 85-103.
91. C. Yi, *et al.*, 2000, Influence of advection on measurements of the net ecosystem-atmosphere exchange of CO_2 from a very tall tower, *Journal of Geophysical Research-Atmospheres*, **105**, 9991-9999.

5. Appendix

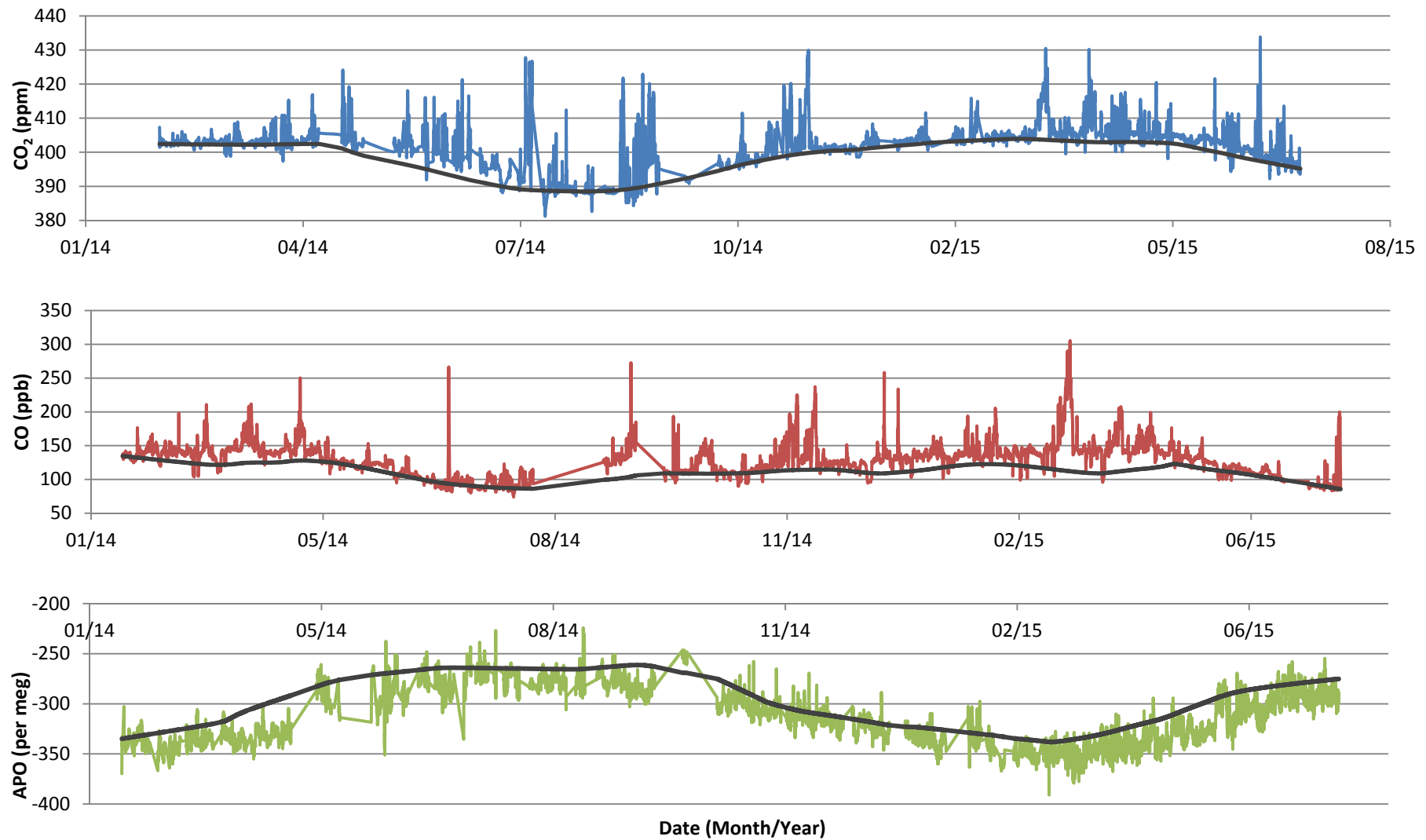


Figure 1 - APO, CO and CO₂ data displayed with the stiff statistical baselines.

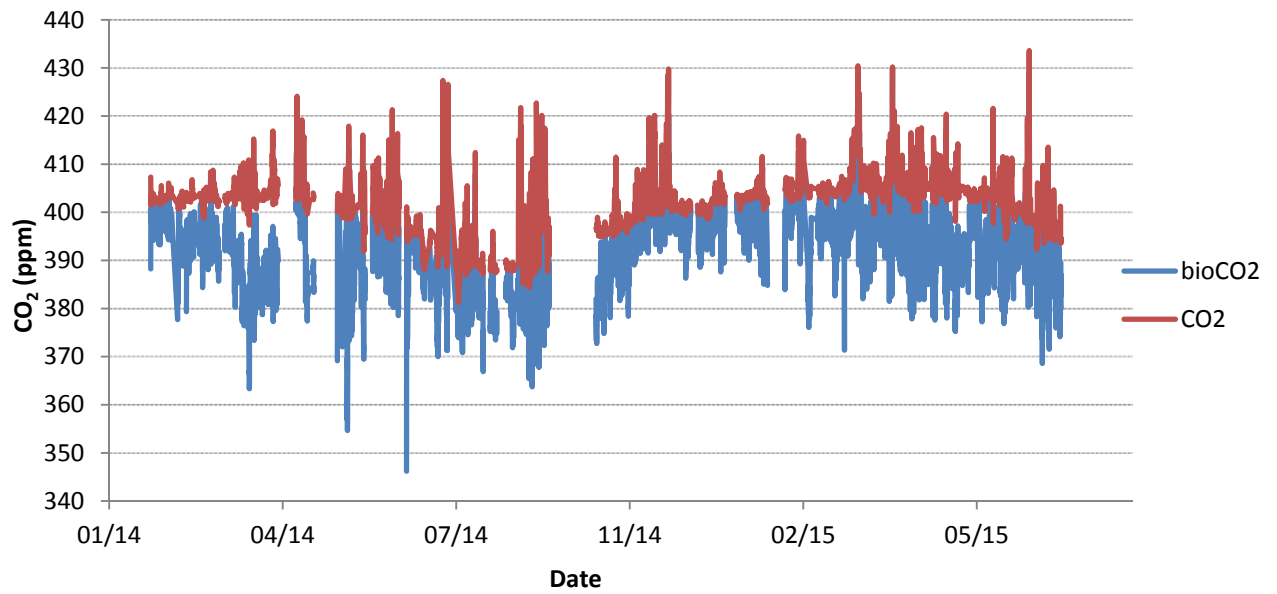


Figure 2 – The obsCO₂ and bioCO₂ for the entire study period. In this plot bioCO₂ was produced using APO as the ffCO₂ proxy along with the stiff statistical baseline.

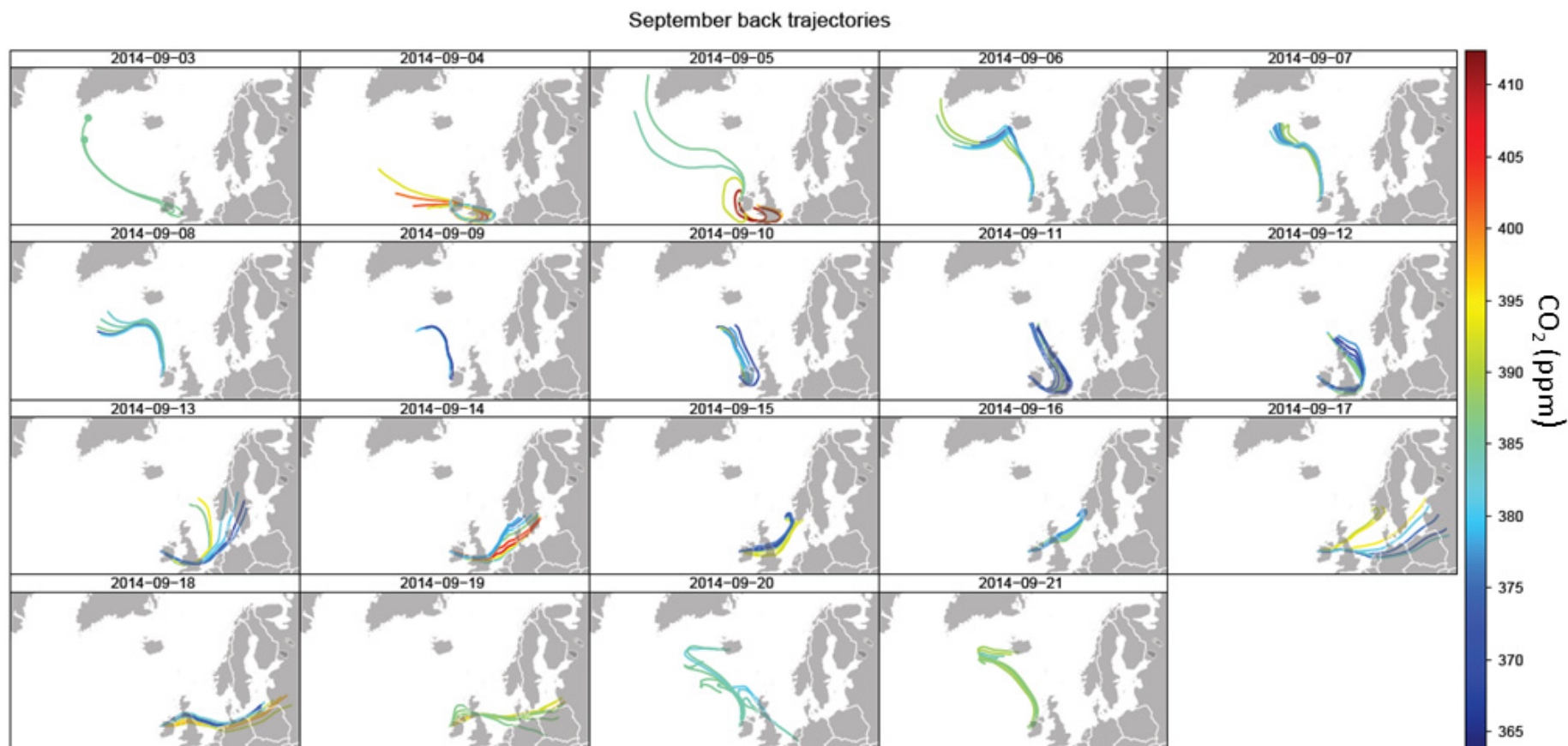


Figure 3 - All of the back trajectories for the September study period are plotted, each spans back 96 hours. The colour of said trajectory represents the bioCO₂ concentration recorded at MHD. The 22nd is not shown as it could not be successfully plotted using R, this may be due to faulty met data or a bug within the program

Table 1 – The dates of back trajectories for each of the clusters

Cluster	Date
C1	4 th (00:00 – 09:00), 10 th (18:00 – 21:00), 11 th , 12 th & 13 th (00:00 – 03:00)
C2	4 th (12:00 – 21:00), 5 th (00:00 – 03:00), 13 th (06:00 – 21:00), 14 th , 15 th (00:00 – 03:00), 17 th (09:00 – 21:00) & 18 th (00:00 – 09:00)
C3	5 th (06:00 – 15:00)
C4	5 th (18:00 – 21:00), 6 th , 7 th , 8 th , 9 th (00:00 – 12:00), 20 th (06:00 – 21:00) & 21 st (00:00 – 15:00)
C5	9 th (15:00 – 21:00), 10 th (00:00 – 15:00) & 21 st (18:00 – 21:00)
C6	15 th (06:00 – 21:00), 16 th , 17 th (00:00 – 06:00), 18 th (12:00 – 21:00), 19 th & 20 th (00:00 – 03:00)