Comments for Editor

The early sections of the Discussion paper have been substantially restructured to follow the suggestions of REF1. The WORD track changes feature shows the changes.

The technical detail in the introductory sections has been reduced:

- Flask Sampling Advantages and Disadvantages is moved to an expanded Supplement (S3).
- The original Section 2 is divided into two sections
    - 2 Background Information on Flask Networks, and
    - 3 Network Inter-comparison
- The scope of the study is more clearly defined at the end of the Introduction – it puts the focus more clearly on variation in the baseline CO_2 records, and their likely relevance to the global carbon budget rather than attempting to revise the carbon budget.

Two new references, and text have been added to address concerns of REF2:

- Adding Bastos et al. (2018) addresses concerns about using only one DVM. Bastos s et al. results are now included in Figure 4 (and prove equally insignificant compared to atmospheric interhemispheric fluxes). Figure 4(d) is added, which now compares the seasonal variability of the contributing processes to CO2 IH differences, and reinforces the basic point of our paper.
- Adding Stavert et al. (2019) contributes to addressing concern about using both mlo-cgo and mlo-spo, as well as supporting the relatively small variability in the ocean exchange.
- The concern about focussing on El Nino and not mentioning La Nina are addressed by including specific references throughout, and by including La Nina anomalies in Figure 9.
- The concern about peak respiration is by-passed by focussing on referring to the CO_2 measurements and their susceptibility to dynamical processes. Details of processes influencing other reservoirs in the global carbon budget are beyond our defined scope.

While these additions do not change the basic premise and conclusions of our study, we found the Referee comments extremely useful in communicating our results to a wider community. The referees are now acknowledged.

Thank you for taking on this difficult task

Roger Francey, 4 September 2019

Response to Referee 1

Interactive comment on "Variability in a four-network composite of atmospheric CO2 differences between three primary baseline sites" by Roger J. Francey et al.

Anonymous Referee #1 Received and published: 3 July 2019

The article by Francey et al describes and analyses 25-year composites of interhemispheric (IH) baseline CO2 differences from NOAA, CSIRO and two independent SIO analysis. They show a good agreement between the 4 monitoring networks and explore the influence of atmospheric dynamics on the CO2 IH gradient with a focus on El Niño periods. The results highlight the importance of IH CO2 transfer parametrization in global carbon cycle models. In general, the paper is scientifically sound and worthy of publication; however, the writing needs some modification and improvement. There is also several typos regarding the results hence the paper needs a careful reading/checking.
We have modified, hopefully improved, and more carefully checked, the writing.

After addressing the comments, the manuscript will be suitable for publication.

General Comments:

1. Currently the introduction sounds like this paper is submitted to AMT and not ACP, and it does not entirely sound like an introduction. For moments it is too techical and it is hard to understand the aim of the work. I recommend re-framing the introduction, to better attract the readers that are interested about this work. I would suggest to include an overview of the key findings of FF16 and FF18 (relevant to the work in this paper), and higlight better what this work aims to add to the previous findings. Parts of referring to results in FF16 and FF18 are indeed included later in the paper; however, the way/order they are represented is making things hard to link together and to see the big picture (and it makes the paper harder to read). Potentially the advantages/disadvantages in the Introduction could be added into a separate section (or combine with Section 2).
A substantial rewrite of the Introduction is in response to these recommendations. Technical discussion is moved to Section 2 or addressed in an expanded Supplement. The Scientific 'big picture' should be clearer, and the different roles of FF16, FF18 and this paper are more clearly specified. The advantages and disadvantages are important when aiming for internal consistency over three decades. However, because of REF1 comments this discussion is now in the Supplement.

2. Usually Introductions end with a paragraph summarizing the aim of the work. I had the feeling that this comes quite late at the end of Section 2. In addition to re-framing the Introduction I would move the last paragraph from Section 2 to the Introduction. This would also partially solve the missing key findings of FF16 and FF18.
The final paragraph of the introduction (Line 68) now defines the scope of the paper.

3. At the moment the clarification/discussion of few things that can potentially lead to confusion for the readers are missing from the paper. It would be good to just summarize somewhere in the paper: 1) the important ENSO periods discussed in the analysis. The reason for this is that I kept thinking that the highlighted 2010 period in Figure 3 had potentially something to do with the strong 2011 La Niña event (although without reading FF16 first). 2) Why La Niña is not discussed at all and how it would impactIH(e.g., LaNiña periods facilitates interhemispheric mixing of trace gases while ElNiño inhibits interhemispheric exchange...). In Section4 the authors wrote "Different responses of IH CO2 to wind indices at different ENSO events, and from non-ENSO periods, are discussed in Section 6." and since ENSO includes both El Niño and La Niña, the impact of La Niña should be at least mentioned somewhere. These things are discussed in FF16 and FF18; however, it would be beneficial to add a sentence or two here also to be easier to track/understand the results.
The concerns about La Nina are mentioned in the Abstract " (line 25)", in lines 63-66 of the Introduction and elsewhere the term ENSO is used to include both El Niño and La Niña events. The ONI in strong La Nina years are now included in Figure 9.

4. Page 1 line 24 and elsewhere in the paper: "5-year relatively ENSO-quiet period" I assume the authors meant "5-year relatively El Niño-quiet period" since 2011 was a strong La Niña period, so ENSO-quiet period is misleading.
Agreed, see point 3 for the revised approach

5. Figure 4 - are there emission anomaly uncertainties that could be added to the Figure and included in the discussion in Section 4.1?
To address this point, we have added a 16 DVM model composite, and seasonal Fossil data to Figure 4 to demonstrate their relative magnitudes. Because all surface-to-air fluxes appear very much smaller than required to significantly alter the baseline CO_2 differences, we have not dwelt on obtaining uncertainties from source data.

Specific Comments:

* Section 2 is quite lengthy and could be simplified. A careful reading to condense some of the text would be useful.

A new section 3 is introduced, and the sampling advantages/disadvantages moved to Supplement S3 in order to break up the lengthy section.

* Page 5 line 183: "e.g. 256 of 300 months have 4 networks" - It seems like the Table shows 268 instead of 256 for mlo-spo, or I misunderstood the Table in which case the authors need to give a better explanation of the Table. Moreover, for mlo-spo with both 268 or 256 the total number of months does not add up to 300. Also regarding the Table, what is the difference between empty boxes and the ones with'-'. If nothing then use consistent marking.

The Table is now updated to 2017 (since SIO spo data became available). The numbers are now consistent, and the main text adjusted accordingly.

* Page 7 line 228: "It is assumed here that flux variations from ocean sources are much smaller than terrestrial" - it would be good to have some references here. Also just mention why anthropogenic (fossil fuel) emissions are not compared here.

This is now addressed in lines 216, lines 315-320, and Figure 8. The seasonal variation in Southern Ocean fluxes (an important part of the global ocean flux) are similar to terrestrial fluxes from the region but much smaller than NH terrestrial flux variation. High precision continuous CO_2 monitoring across the Southern Ocean (Stavert et al., 2019, and personal communication) confirm relatively small variation. Fossil emission seasonal variations are included line 322 and in the new Figure 4(d).

* Page 9 line 334: It would be good to state why did the authors choose the ONI index as the ENSO index instead of the other indices. It would be also interesting to see if other ENSO indices show the same results (but not strictly necessary to include in this paper if it is too time consuming).

We have re-examined the ONI data and they give essentially the same result as Nino3 and Nin3.4 indices, as stated in line 331.

*Page9 line337: "there are no significant ONI"-well relative to ElNiño but not relative to La Niña, so the statement is a little bit misleading.

See reworded lines 3333-336.

Technical Comments:

*Page2 1st and 2nd paragraph: It feels like a weird jump between the two paragraphs, a 'linking' sentence between the two would be good.

A linking sentence is provided, lines 33-36.

* Page 5 line 166, Page 8 line 308, Page 9 line 334: however -> ;however, * Page 6 line 200: & -> and

Fixed, thanks

* Page 8 line 271: FF18.These -> missing space before These

Fixed, thanks

* Figure 5: ITCZ -> Inter Tropical Convergence Zone, the abbreviation was not defined

Fixed, thanks

* Page 9 line 314: "see Discussion, Section 5" -> see Discussion, Section 6 maybe?

Fixed, thanks

* Page 9 line 339 ppm.(PgC) -> missing space before (PgC)

Fixed, thanks

* Page 10 line 258-359: why is this a separate paragraph?

Fixed, thanks

* northern hemisphere is somewhere capitalized and somewhere not in the text * NH abbreviation is not defined

Fixed, thanks

* Figure 3 inconsistency between figure and caption: (a) mlo-cgo, (b) mlo-spo -> (a) mlo-spo, (b) mlo-cgo

Fixed, thanks

* Figure 7 and 8 please add x axis label

Fixed, thanks

Response to Referee 2

Interactive comment on "Variability in a four-network composite of atmospheric CO2 differences between three primary baseline sites" by Roger J. Francey et al.

Anonymous Referee #2 Received and published: 16 July 2019

Summary: Francey et al. present an analysis of monthly data for three measurement stations (MLO, CGO and SPO) to investigate the interhemispheric CO2 difference (IH CO2, defined as concentration difference of MLO-CGO or MLO-SPO) variations over the last 25 years. After comparing the different data sets, the potential biogeochemical drivers for the short-term, seasonal and inter-annual variability of IH CO2 is discussed. The manuscript highlights the ability of different atmospheric transport indices to explain these IH CO2 variations. The manuscript is well-structured and contains important and refined ideas that will be useful in guiding the improvement of GCMs commonly used in global carbon cycle modelling. Its topic is well suited for ACP and surely of interest to the wider scientific community. However, two general comments and a few minor comments should be addressed before publication.

General comments: One key argument of this paper is that the variations of IH CO2 cannot be explained by the net uptake of CO2 in the NH terrestrial biosphere. Unfortunately, this assumption is based only on a single DGVM, which are known to be highly uncertain. A comparison of CABLE and other models can be found here: https://journals.ametsoc.org/doi/pdf/10.1175/2008JCLI2378.1. Multiple DGVMs or an ensemble mean of CMIP LSMs would be more robust here.

A 16-DGVM ensemble of extra-tropical NBP from Bastos et al. 2017 is now included (lines222-225), including representation in Figure 4. It confirms the surface flux anomalies are too small to account for IHΔCO$_2$ changes.

When considering fluxes from optimized emission products (e.g. CarbonTracker-EU) a significantly larger share of IH CO2 could be explained by the NH terrestrial biosphere, although likely not contradicting the finding that IH CO2 is strongly influenced by transport processes.

"We do not discuss air-surface fluxes derived from CO$_2$ data that are less spatially representative, and/or rely on atmospheric transport modelling. The latter introduce additional model degrees of freedom and potentially overestimate terrestrial variability if the variability in atmospheric IH transport is not adequately captured" (lines 224-227).

The manuscript does not use consistent language around the main driver of NH CO2 enhancements. The claim that respiration is maximal in early parts (FEB-APR MAY-JUL) seems unlikely and needs to be supported. However, later the authors refer to NBP and "terrestrial emissions" (which could likely peak in different seasons than autotrophic & heterotrophic respiration. As this is a key issue a clarification would be most helpful.

This is a valid criticism. The scope of this paper is to explore and seek explanation for variations in the baseline CO2 records. While there are clear implications for the global carbon budget, lack of advanced knowledge in every individual component of the budget limits a more comprehensive approach. Reducing uncertainty in the critical atmospheric component of the global budget is the purpose of this study. This is clarified in Lines 68-73 of the rewritten Introduction an in responses to *Specific comment Lines 31, 131, 172, 190*.

Specific comments:

Line 31 and Line 44: Why is the proposed method only compared to growth rate studies here and not to 4D-VAR and ensemble kalman filter/smoother data assimilation systems that have been shown to reproduce global scale fluxes with fairly reasonable performance?
The Introduction rewrite aims to clarify this. See also lines 222-225 generated in response to General comments.

Line 55: Figure 1 seems to suggest that with-in network errors do not cancel our. No consistent offset between the different measurement programs was found for all sites. However, to assume that the difference is small compared to the IH CO2 signal seems perfectly reasonable.
We make this assumption.

Line 60: Multi-species observations are mentioned here, but not discussed in this manuscript. Examples of relevant species and isotopes should be given (or paragraph removed).

Multi-species studies, outside the current scope, have the potential to further inform this debate. They are briefly mentioned at line 63 and in the expanded Supplementary information.

Line 63+: The authors do not mention the added value of quasi-continuous data from in-situ observations. If they were available (or used here), issues relating to the impact of different sampling dates and sampling frequencies of the different flask programs could assessed more quantitatively.

The advantages and disadvantages are relevant when aiming for internal consistency over three decades. Challenges to achieving consistent site differences are much greater with the conventional NDIR analysers used over most of this period. This discussion is now focussed in the new Supplement. In relation to spatial differences, NDIR in situ analyses have had calibration limitations due to short lifetime of reference and calibration gases. Flasks analyses in a central laboratory have largely avoided these limitations over decadal timeframes.

Line 122: Unclear what kind of information is available or referenced in Keeling 1998. It would be great to have more details.

It is a brief reflection of the difficulties in maintaining government support for monitoring. It is best articulated in the original publication.

Line 131. Please clarify: the maximum of 7-10ppm of IH CO_2 occurs when most exchange occurs between NH and SH? Seems counter-intuitive or do you talk about the specific gradients of two sites here?

Yes, poorly worded. A CO_2 partial pressure difference between hemispheres is a prerequisite for net IH exchange. See lines 120-1.

Line 133: Do the offsets really "cancel" here? (see Line 55).

The words used were "largely cancel", which is the case. (It is supported in Figure 4 by the ΔIHCO_2 response to the 1997 GFED anomaly).

Line 155: Why are those definitions of seasons used here and not the more common meteorological or astronomical definitions? Please consider clarifying.

The reasons are stated in lines 141-143: "the particular 3-month seasonal selection distinguishes periods of relatively stable partial pressure differences between hemispheres and the selected seasons also distinguish eddy and mean IH transport mechanisms (FF18)"

Line 167: Why was the decision made not corrected or flag the flask data but "average out" potential outliers by making a composite time series?

This sentence generally describes selection made by each laboratory prior to publication of the monthly averages used here. In the subsequent compositing this data clear statistical outliers are revealed of unknown origin. In the composite these are suppressed by averaging. See lines 115-118.

Line 172: Please provide a reference on the assumption that maximum respiration from NH forests occurs in the seasons in FEB-APR and MAY-JUL? This seems fairly unlikely, as especially FEB-APR is still cold in most of the boreal forest regions in the NH and both autotrophic and heterotrophic respiration typically increase later in the year e.g. in (late) summer. Or please clarify if this refers to net biome productivity or net ecosystem exchange or "terrestrial emissions" (mentioned later in the manuscript).

See response at the end of General Comments above. The wording at line 157 now merely states the widely accepted general reason for the NH CO_2 seasonality.

Line 183: The data in table 1 seems inconsistent with the example here. In general, 1992-2016 is mentioned, but later 1992-2017 is shown. (Figure in supplement has 288 months of maximum data, main text figure refers to 300 months maximum).

The Table is now updated to 2017 (since SIO spo data became available). The numbers are now consistent, and the main text adjusted accordingly.

Line 190: Please clarify "generally attributed to NH forest..." this manuscript argues (in later sections) that 1.) the IH CO_2 gradient is dominated by transport and that 2.) the variability in the IH CO_2 is driven by the MLO time series. Is this consistent?

The seasonality in mlo CO_2 is generally attributed to NH forests, and SH seasonality is small by comparison.

However, the anomalies (mean seasonality subtracted) correspond primarily to anomalies in IH transport indices.

Line 199: Why is 2017 now included (see comment L183)
See reply to comment L183.

Line 214: Some technical detail on how the fit was done would be most useful to the reader here, did you perform a sensitivity analysis for different types of splines?
Spline polylines, available in commercial plotting software, link peaks and dips to visually aid discussion.

Line220: Please clarify: the assumption that the global atmospheric CO2 budget is not driven by IH transport but by exchanges to other compartments (biosphere, oceans, etc.) does not have to rely on assuming that transport is correct. The total global atmospheric CO2 budget is the same no matter if CO2 is in the NH or SH. I assume the authors want to argue that the NH versus SH budget might be significantly wrong when not accounting for IH exchange or maybe that the sites used (e.g. MLO) reflect more signals than just emissions and sinks in NH?
Yes. Wording in Section 5 is adjusted to make this clearer.

Line 235+: Why was the CABLE model used. DGVMs are inherently uncertain, so using multiple DGVMs or an ensemble estimate would give a better representation of the range of estimates of NH terrestrial fluxes. Optimized emission products, e.g. CarbonTracker-Europe report significantly higher NH terrestrial uptake (same order of magnitude as IH processes discussed later in the manuscript and in Figure 4). Available for download at: http://www.carbontracker.eu/fluxtimeseries.php [See also general comment]
These issues are addressed in the response to General comments, above.

Line253: The qualitative study on the potential impact of CO2 emissions on global and hemispheric CO2 concentrations seems very instructive. However, a more qualitative estimate using a GCM could be beneficial here.
Maybe, if the GCM has sufficient resolution and upper equatorial tropical parameterisation? It is outside the scope of this study.

Line 286: The authors raise an interesting point here: the MLO-SPO and MLO-CGO data is "effectively identical". Why was SPO data included in this study? A separate study of CGO-SPO could have been more enlightening if it could address the question of CO2 uptake in the Southern Oceans.
This point is specifically addressed in the new introduction. CO_2 data quality (measurement and spatial representation) is considered a key issue in this paper, in which the composite plays a central role. The cgo-spo comparison is the best independent evidence for both factors. Note: The Southern Ocean uptake is informed by other records (including in ultra-high precision in situ monitoring at cgo and Macquarie Island plus other Antarctic sites). This is outside the scope of this paper and is the subject of on-going investigation by others (e.g. Stavert et al.). And we speculate that other factors such as seasonality in FF emissions may be involved, see line 315-322.

Line 345: What are the uncertainties of the estimated annual concentration trends and are they statistically significant (and different)?
Statistical uncertainties in trends, are now included throughout the paper. We also provide an estimate of uncertainty in the annual average data of Figure 9, Line 326.

Line 368: Here, the manuscript refers to "terrestrial emissions" is this equal to F(ffco2)+NBP?
Yes, text amended, Line 367.

Line 404: How would the IHCO2 impact the growth rate of CO2 derived at other long term reference sites in the NH. Like Barrow, US and Alert, CA? Could analyzing data from those sites help to further separate the impact of IH mixing versus NH fluxes?
Possibly, but spatial representation and quality of data are limiting factors.

TRACK CHANGES

Variability in a four-network composite of atmospheric CO$_2$ differences between three primary baseline sites

Roger J. Francey, Jorgen S. Frederiksen, L. Paul Steele and Ray L. Langenfelds

CSIRO Oceans and Atmosphere, Aspendale, Victoria, AUSTRALIA
Correspondence to: Roger J. Francey (roger.francey@csiro.au)

Abstract. Spatial differences in the monthly baseline CO$_2$ since 1992 from Mauna Loa, (mlo, 19.5°N, 155.6°W, 3379m), Cape Grim (cgo, 40.7°S, 144.7°E, 94m) and South Pole (spo, 90°S, 2810m), are examined for consistency between four monitoring networks. For each site pair, a composite based on the average of NOAA, CSIRO and two independent SIO analysis methods is presented. Averages of the monthly standard deviations are 0.25, 0.23 and 0.16 ppm for mlo-cgo, mlo-spo and cgo-spo respectively. This high degree of consistency and near-monthly temporal differentiation (compared to CO$_2$ growth rates) provides an opportunity to use the composite differences for verification of global carbon cycle model simulations.

Interhemispheric CO$_2$ variation is predominantly imparted by the mlo data. The peaks and dips of the seasonal variation in interhemispheric difference act largely independently. The peaks mainly occur in May, near the peak of Northern Hemisphere terrestrial photosynthesis/~~terrestrial~~ respiration cycle. ~~Boreal spring~~Feb-Apr is when interhemispheric exchange via eddy processes dominates, with increasing contributions from mean transport via the Hadley circulation into boreal summer (May-Jul). The dips occur in September, when the CO$_2$ partial pressure difference is near zero. The cross-equatorial flux variation is large and sufficient to significantly influence short-term Northern Hemisphere growth rate variations. However, surface-air terrestrial flux anomalies would need to be up to an order of magnitude larger than found to explain the peak and dip CO$_2$ difference variations. ~~, just after the peak in the mean interhemispheric exchange via the Hadley circulation. Surface-air terrestrial flux anomalies would need to be up to an order of magnitude larger than found in order to explain the peak and dip CO$_2$ variations (large enough to significantly influence short-term northern hemisphere growth rate variations).~~

~~F~~Recent ~~f~~eatures ~~in~~ throughout the composite records, inconsistent in timing and amplitude with air-surface fluxes, are largely consistent with interhemispheric transport variations. These include greater variability prior to 2010 compared to the remarkable stability in annual CO$_2$ inter-hemispheric difference in the 5-year relatively El Niño-quiet period 2010-2014 (despite a strong La Nina in 2011), and the 2017 recovery in the CO$_2$ interhemispheric gradient from the unprecedented El Niño event in 2015-16. ~~These include the remarkable stability in annual CO$_2$ inter-hemispheric difference in the 5-year relatively ENSO-quiet period 2010-2014, and the 2017 recovery in the CO$_2$ interhemispheric gradient from the unprecedented El Niño event in 2015-16.~~

1   Introduction

Atmospheric CO_2 measurements are normally introduced into global carbon budgets as a "global growth rate ... based on the average of multiple stations selected from the marine boundary layer sites with well mixed background air .., after fitting each station with a smoothed curve as a function of time, and averaging by latitude band …" (Le Quéré et al., 2018). This approach encourages sampling at multiple locations to seek atmospheric confirmation of national/continental emission changes. Particularly in the Northern Hemisphere (NH), with more complicated geography and atmospheric circulation, the influence of continental emissions on marine boundary layer air can vary widely between sites.

A clearer indication of the global impact of regional emissions comes from sites demonstrating maximum spatial representation. In this case, global significance of biogeochemical CO_2 exchanges between the surface will be informed by their impact on validated baseline data with the least continental influence. Such baseline data is more directly relevant to changes in global ocean acidification and climate change, but places heightened demands on sampling criteria and calibration.

Sites selected to maximise spatial representation in their respective hemispheres, Mauna Loa, (mlo, 19.5°N, 155.6°W, 3379m) and South Pole (spo, 90°S, 2810m) also have the longest (multi-decadal) coherent trace gas monitoring data, based on flask sampling (Supplement S1). At these sites, and at Cape Grim (cgo, 40.7°S, 144.7°E, 94m) since 1991, co-sampled baseline air has been analysed at three different laboratories, using four different methodologies summarized in Section 2.

To account for any persisting artefacts in the co-sampled data we examine, for each method, inter-site differences in the published monthly baseline data from the three sites. The standard deviation in the average of the co-sampled differences provides a practical uncertainty estimate. A key advantage compared to the growth rate approach, is that assumptions inherent in the growth rate smoothing (where for example 22-month smoothing is used to separate interannual and seasonal variations) are avoided so that in this study near-monthly effective time resolution is achieved.

The inter-site difference approach was used by Francey and Frederiksen, 2016 (FF16) to conclude that the suppression of the normal eddy component of interhemispheric (IH) CO_2 exchange in Feb-Apr 2010 contributed to the unprecedented 0.8 ppm step in the IH difference between 2009 and 2010. The dynamical anomaly was associated with a moderately strong El Niño leading to a NH build-up of CO_2 in 2010. FF16 supplementary information demonstrated a failure of atmospheric transport models of the carbon cycle to simulate the step.

The 2015/16 El Niño was stronger and has also been associated with unprecedented behaviour in the global carbon cycle (elsewhere attributed to the terrestrial biosphere anomalies, e.g. Yue et al., 2017). However, Frederiksen and Francey, 2018 (FF18), argued that the unprecedented strength in the Hadley circulation increased IH exchange (reduced IH CO_2 difference) late in 2016, overwhelming the earlier reduced eddy exchange linked to the strong 2015/16 El Niño. They also indicated dynamical contributions to IH CO_2 during both El Niño and La Niña periods (e.g. FF16 Figure 5, and FF18 section 6.2, on multi-species IH differences). While ENSO events are expected to impact on surface biology, it is also clear that they also influence atmospheric IH CO_2 fluxes. The timing of the dynamical events suggests an alternate explanation for the CO_2 behaviour discussed by Yue et al., if it can be demonstrated that IH CO_2 fluxes at the time exceed their postulated air-surface terrestrial fluxes.

The scope of this paper includes: a) reduction of measurement uncertainties in IH CO_2 difference using a three decade composite of published CO_2 measurement results (distinguished by maximum spatial representation and by well-documented sampling and measurement quality), and b) demonstration of the potential uses of the composite CO_2 record by comparing anomalies in the magnitude and phasing of composite IH CO_2 variations with those in air-surface exchange model outputs, as well as in dynamics indices representing atmospheric IH exchange.

2 Background information on flask networks

A historic overview of CO_2 IH difference data is provided in Supplement S1.

By 1958 C.D. Keeling had identified mlo and spo as optimum sites to obtain background CO_2 in the respective hemispheres and by the 1970s was obtaining a regular monthly supply of air admitted to 5L evacuated glass flasks from both sites (SIO1: Keeling C. D. et al., 2001). Since 1992, there are CO_2 measurements as a by-product of a global network focussed on O_2/N_2 ratios in baseline air (SIO2: Keeling R. F. and Schertz, 1992); this program uses 5L glass flasks flushed and filled to ambient pressure, with cryogenically dried air. While there is commonality regarding calibration, in the context of spatial differences the SIO networks can be considered independent.

NOAA began sampling from all three sites, mlo, spo, cgo, (as part of a much larger network) from 1984, using a variety of flask and filling methods. From around 1992 the current system of Peltier-dried air in pressurized 2.5 L flasks (Tans et al., 1992, Conway et al., 1994, Dlugokencky et al., 2014) was phased in. NOAA has maintained the World Meteorological Organization (WMO) Central CO_2 Calibration Laboratory since 1996 (a role previously carried out by SIO). The NOAA atmospheric sampling is generally more frequent (typically 8-10 flasks per month) than is the case for the SIO or CSIRO programs (except for the CSIRO cgo program); however, the size and sampling frequency in the NOAA network amplifies calibration challenges due to shorter lifetimes of reference and calibration standards.

Both NOAA and SIO use non-dispersive infra-red analysers (NDIR) for CO_2 measurement. (CSIRO flask sampling at cgo, spo and mlo in the early 1980s used NDIR for analysis of chemically dried air, pressurized into 5L glass flasks.) However, analyses here are restricted to CSIRO's measurements from 1992 using chemically-dried, pressurized air in 0.5L glass flasks, but with retention of 5L flasks at spo (Francey et al., 1996). Gas chromatography with flame ionisation detection (GC/FID) was introduced to measure CO_2 in flasks, a technique providing a more linear response than NDIR (Supplement S2). Hourly radon measurements at Cape Grim (Chambers et al. 2016) were introduced around this time. Air mass history is further informed by a decade of vertical profiling (Langenfelds et al., 2003; Pak et al., 1996), back trajectory analysis, and other tracers (e.g. Dunse et al., 2001), demonstrating that selected cgo data can achieve a degree of spatial representation matching, or sometime exceeding, that at the more remote high-altitude sites at mlo and spo.

Apart from longevity, the flask records offer other advantages over in-situ monitoring, but are more susceptibile to some unfavorable factors as are discussed in Supplement S3. The challenges of maintaining high quality over decades in any one monitoring program are many. They include external factors, acknowledged but not pursed here, such as high turnover of skilled staff particularly at remote air sampling sites or changes in institutional strategic and economic priorities. The latter are well described by Keeling (1998), with CSIRO sharing similar institutional experiences.

3 Interhemispheric (IH) baseline CO$_2$ differences have been linked to intermittency in seasonal exchange between hemispheres by Francey and Frederiksen, 2016 (FF16) and Frederiksen and Francey, 2018 (FF18). The dynamical processes described in these studies, if not adequately captured in global carbon cycle model transport, will compromise conventional approaches that assume sub-annual CO$_2$ variation is primarily the result of terrestrial exchange (e.g. Yue et al., 2017, Rödenbeck et al., 2018).

FF16 and FF18 focussed on Commonwealth Scientific and Industrial Research Organisation (CSIRO) CO$_2$ data. However, co-sampling over 25 years of background CO$_2$ has occurred at three iconic remote sites by four well-established global monitoring flask sampling networks, two from Scripps Institution of Oceanography (SIO1, SIO2), one from the National Oceanic and Atmospheric Administration (NOAA) and one from CSIRO. Consistency in a composite of within-network spatial differences provide an improved opportunity to examine inherent assumptions in the statistical treatment of atmospheric data in global carbon budgets.

Compared to CO$_2$ information used in typical growth rate studies, the 25-year composites of within-network spatial differences measured by the flask networks provide additional insights, summarised here:

The effective time resolution can be short, circumventing ambiguities between seasonality and inter-annual variability, and 6-9 month end-effects inherent in conventional growth rate analyses. The latter are generally obtained by harmonic filtering to define seasonality, and 22 month smoothing to define inter-annual behaviour (e.g. Thoning et al., 1989). The monthly average spatial differences used here have higher resolution, being obtained from 80-day smoothed records that represent the available flask measurements.

Since flask samples from each network are analysed in a central laboratory, bias associated with calibration of an internal calibration scale relative to the international CO$_2$ WMO mole fraction scale (Zhou and Tans, 2006), and similar bias when relating a reference gas to the internal scale, generally cancel for within-network site differences.

The IH CO$_2$ differences suppress the influence of equatorial surface exchanges that are uplifted to an altitude where they can mix into both hemispheres. This improves sensitivity to cross-equatorial atmospheric fluxes that occur in a region where transport is less well-defined than at higher latitudes (e.g. Lintner et al., 2004).

Multi-species (e.g. other long-lived trace gases and their isotopes) with different biogeochemistry but identical time of collection are often available (particularly from pressurized flasks). A multi-species approach was briefly assessed in FF16 and FF18 and is the subject of further studies.

Disadvantages of flask compared to *in situ* CO$_2$ sampling include:

CO$_2$ artefacts related to extended storage of air in flasks, mainly affecting sampling from remote sites with annual resupply (in this study spo). Flask size, permeation through elastomer seals in pressurized flasks (Sturm et al., 2004), and inadequate surface preconditioning to limit CO$_2$ adsorption on surfaces, are possible contributing factors. In this study, factors such as the use of larger and/or unpressurized flasks at spo, and corrections informed by co-sampling with *in situ* analysers (e.g. Stavert et al., 2019), help address this concern.

The brevity of sampling compared to *in situ* measurement. Monthly average concentrations from each site comprise the average of 1-10 or more flask samples

3 Network intercomparison

2 Background information on flask networks

By 1958 C.D. Keeling had identified mlo and spo as optimum sites to obtain background CO_2 in the respective hemispheres and by the 1970s was obtaining a regular monthly supply of air admitted to 5L evacuated glass flasks from both sites (SIO1: Keeling C. D. et al., 2001). Since 1992, there are CO_2 measurements as a by-product of a global network focussed on O_2/N_2 ratios in baseline air (SIO2: Keeling R. F. and Schertz, 1992); this program uses 5L glass flasks, with air cryogenically dried, flushed and filled to ambient pressure. While there is commonality with regard to calibration, in the context of spatial differences the SIO networks can be considered independent.

NOAA began sampling from all three sites (as part of a much larger network) from 1984, using a variety of flask and filling methods. From around 1992 the current system of Peltier-dried air in pressurized 2.5 L flasks (Tans et al., 1992, Conway et al., 1994, Dlugokencky et al., 2014) was phased in. They have maintained the WMO Central CO_2 Calibration Laboratory since 1996 (a role previously carried out by SIO). The NOAA atmospheric sampling is generally more frequent (typically 8-10 flasks per month) than is the case for SIO or CSIRO programs (except for the CSIRO cgo program). However, the size and sampling frequency in the NOAA network amplifies calibration challenges (e.g. due to shorter lifetimes of reference and calibration standards).

Both NOAA and SIO use non-dispersive infra-red analysers (NDIR) for CO_2 measurement.

CSIRO flask sampling at cgo, spo and mlo in the early 1980s used chemically dried air, pressurized into 5L glass flasks, using NDIR for analysis; however, analyses here are restricted to CSIRO's more comprehensive measurements from 1992 using chemically dried, pressurized air in 0.5L glass flasks, but with retention of 5L flasks at spo (Francey et al., 1996). In the early 1990s the use of gas chromatography with flame ionisation detection (GC/FID) was exploited to monitor CO_2 in flasks at CSIRO. The GC/FID technique used provides a significantly more linear response for CO_2 than NDIR (see Supplement Figure S1) and has required much smaller samples than most other equipment employed consistently over the 25 years; both factors contribute to calibration

~~integrity. Counterbalancing the fact that smaller sample flasks imply enhanced susceptibility to flask storage effects, the small sample requirement has permitted pioneering, long term "same air" measurement inter-comparisons between NOAA and CSIRO on NOAA cgo samples (Masarie et al., 2001; see Supplement Figure S3). Hourly radon measurements at Cape Grim (Chambers et al. 2016) were also introduced around this time. Cape Grim sampling is further informed by a decade of vertical profiling (Langenfelds, 2003; Pak, 1996), back trajectory analysis, and other tracers (e.g. Dunse et al., 2001), demonstrating that selected cgo data can achieve a degree of spatial representativeness that matches, or sometime exceeds, that at the more remote high altitude sites mlo and spo.~~

~~Note: The challenges of maintaining high quality over 25 years in any one monitoring program are many. They include external factors, acknowledged but not pursed here, such as high turnover of skilled staff particularly at remote monitoring sites or changes in institutional strategic and economic priorities. The latter are well described by Keeling (1998), with CSIRO sharing similar institutional experiences.~~

NOAA, which has the most extensive global network (and since 1996 has also operated the WMO CO_2 Central Calibration laboratory) is selected as the reference for an initial inter-network comparison. For each of the three baseline sites, Figure 1 shows systematic behaviour in the SIO1, SIO2 and CSIRO monthly CO_2 differences from NOAA. Five-month running means aid discussion. ~~Before combining data from different networks, systematic differences between programs and sites are examined in Figure 1. It shows monthly differences from NOAA data of SIO1, SIO2 and CSIRO data, for each of the three baseline sites. Five month running means aid discussion. NOAA, because it has the most extensive global network, and since 1996 has also operated the WMO CO_2 Central Calibration laboratory, is selected as the reference.~~ The CO_2 mixing ratios used here are referred to in the commonly used units of parts per million (ppm) rather than the more strictly correct term of μmole of CO_2 per mole of dry air. Note that data independently flagged for sampling or measurement anomalies are rejected by individual laboratories prior to publication as monthly averages. Typically, a small number of gross outliers in individual flask data (e.g. in flask-pair differences) are also rejected prior to publication.

In Figure 1, there is clear evidence of systematic difference in mean offsets, seasonality, and between sites within one network. In the context of inter-hemispheric exchange, the typical 0.5 ppm range of variation remains relatively small compared to the 7-10 ppm maximum CO_2 interhemispheric difference (IH ΔCO_2). Net IH exchange is proportional to IH partial pressure difference. ~~IH CO_2 (when most exchange occurs). Also, when calculating within-network IH differences consistent calibration offsets at all three sites will largely cancel.~~

- Between 1991-1993, there is a marked inconsistency between NOAA mlo-spo and mlo-cgo, particularly in seasonal amplitude; CSIRO has comparable measurements that are more consistent (Supplement S4). This is a reason for caution when interpreting the data in this period.
- In the post-1996 statistics the SIO1 offsets from NOAA behave similarly for mlo and spo. This is not the case for SIO2, which has similar offsets at mlo (-0.18 ppm) and cgo (-0.19 ppm), but not at spo (-0.04 ppm), or for CSIRO (mlo: -0.08 ppm, cgo: +0.01 ppm, spo: +0.13 ppm).
- CSIRO records at cgo exhibit the smallest offset and scatter relative to NOAA (± 0.08 ppm) while SIO2 mlo data exhibit the largest scatter (± 0.37 ppm).

- Remnant seasonality is still evident in the CSIRO cgo differences from NOAA. While a small effect, the CSIRO GC/FID near-linear response for CO_2 means results are not so sensitive to differences between sample and reference CO_2. This advantage is reinforced in the CSIRO SH data since reference gases use recent SH baseline air. This is generally not the case for non-linear NDIR measurement and particularly in the NH if relatively short-lived reference gases sourced in the NH have a less-than-optimum match with ambient CO_2 from a site.

The means and bracketed standard deviations in the differences from NOAA that are included in Figure 1 are calculated for data from the years 1996-2016; prior to 1996, variability is generally greater and corresponds to a period of change and development in all three laboratories. For example, there is a marked inconsistency between NOAA mlo-spo and mlo-cgo in 1991-1993, particularly in seasonal amplitude; CSIRO has comparable measurements that are more consistent (see Supplement Figure S3). This is a further reason for caution with the data in this period.

In the post-1996 statistics the SIO1 offsets and scatter from NOAA behave similarly for mlo and spo. This is not the case for SIO2, which has similar offsets at mlo (-0.18 ppm) and cgo (-0.19 ppm), but not at spo (-0.04 ppm), or for CSIRO (mlo: -0.08, cgo: +0.01, spo: +0.13 ppm). SIO2 mlo data exhibit the largest scatter (±0.37 ppm). The CSIRO records at cgo exhibit the smallest offset and scatter relative to NOAA (+0.01 ± 0.08 ppm), but remnant seasonality is still evident. A possible consideration here is that the CSIRO GC/FID near-linear response for CO_2 means that results are not sensitive to differences between sample and reference CO_2, a factor reinforced in the SH if reference gases use recent SH baseline air. This is generally not the case for non-linear NDIR measurement and particularly in the NH if relatively short-lived reference gases sourced in the NH do not match ambient CO_2 from a site.

Systematic differences due to sampling and measurement methodology are likely the result of a combination of the linearity of instrument response and flask storage effects (particularly at spo and particularly in CSIRO spo data with the sparsest sample density), that is likely to be partly addressed with a comprehensive review of metadata. This is outside the scope of this study so we rely on the ensemble of average differences to moderate their influence.

While Figure 1 reveals some un-resolved systematic differences between data sets, Figure 2 emphasizes that they are generally small compared to the IH partial pressure differences that are a pre-requisite for IH net exchange. Figure 2 provides an overview of the impact of measurement bias on spatial differences. Data from each method are presented as 3-month seasonal averages in order to minimize potential influences related to network sample-frequency (by ensuring an adequate number of individual flask samples per seasonperiod). As well, the particular 3-month seasonal selection distinguishes periods of relatively stable partial pressure differences between hemispheres and the selected seasons also distinguish eddy and mean IH transport mechanisms (FF18).

Figure 2 demonstrates the considerable coherence between data sets.

- For the most part, and particularly in the Aug-Oct season when IH ΔCO_2 IH CO_2 difference (IH ΔCO_2) is at a minimum, there is a high level of consistency in the year-to-year variation in seasonal spatial differences from each network.
- There are relatively few examples of one record differing markedly from the others; when it occurs, it is often for reasons evident in Figure 1. For example, in Figure 2 NOAA cgo-spo appears low in 1992/1993;

CSIRO mlo-spo shows negative outliers in May-Jul 2009 and Nov-Jan 2002, but not at mlo-cgo. SIO2 outliers in 2002 and 2006 exhibit similar characteristics; positive outliers e.g. SIO2 Nov-Jan 2016 suggests a cgo problem. In Feb-Apr 2005 NOAA data indicate a possible mlo problem; however, this is also when the 'volatility' of the records (and in IH transport) is large, so it is conceivable that different flask sampling numbers and times could contribute to lower values by both SIO and CSIRO. ~~Closer inspection of individual flask metadata may resolve these infrequent anomalies, but for the present, composite averaging is relied on to moderate their influence.~~

- The largest IH ΔCO_2 variability is recorded in Feb-Apr and in May-Jul, both seasons having near-equally large IH differences. These seasons correspond to maximum respiration from NH forests. Feb-Apr is also when IH exchange by eddy processes is most influential (FF16), whereas mean transport via the Hadley circulation is the main dynamical influence in May-Jul (FF18).

  - Systematic differences due to sampling and measurement methodology can possibly arise from factors such as the linearity of instrument response, flask storage effects or undetected entrainment of laboratory air. Records with the sparsest sample density (e.g.at spo and particularly in CSIRO spo data) may be more susceptible to undetected anomalies. Closer inspection of individual flask metadata, or of the less extensive *in situ* monitoring, may resolve some of these infrequent anomalies, but for the present, composite averaging of the flask data is relied on to moderate their influence.

  -

~~4 FF18 provided correlations of mlo-cgo with wind indices representing both eddy and mean IH transport (using CSIRO data only), emphasizing temporal covariation. FF16 examined the relationship between eddy transport indices and IH CO₂ differences for CSIRO mlo-cgo and SIO1 mlo-spo data. Here we individually examine major anomalies in the 25-year composite records with more emphasis on the magnitude and timing of influences that might be attributable to IH transport variation.~~

4 Composite records of baseline station spatial differences

For each of mlo-cgo, mlo-spo and cgo-spo monthly CO_2 differences, Table 1 shows the number of months between 1992 and 2016 contributing to a composite value, arranged in columns indicating the number of contributing networks; e.g. 256 of 300 months have 4 networks contributing to mlo-spo, while 277/300 months have three contributing networks at mlo-cgo.

The percentage of missing months for each network, and scatter in the composite differences for different historic periods is tabulated in the Supplement ~~Tables~~ S1~~, S2. Data analysed below extend to 2017.~~

The monthly composite CO_2 differences are shown in Figure 3 (and tabulated in ~~our~~ Supplement Table ~~S3~~S5). The small ~~e~~Error bars represent the ensemble standard deviation (The one ~~exce~~ption is ~~for the~~ cgo-spo ~~case,~~in Feb 2009, with only the NOAA network contributing; it is arbitrarily assigned 100% uncertainty and appears as an outlier in Figure 3c). The seasonality at mlo, generally attributed to NH forest photosynthesis/respiration cycles, is the dominant variation in IH ΔCO_2~~the IH CO₂ differences~~. The composite errors are small ~~by comparison~~compared to seasonal amplitudes. Average ~~s~~Standard deviations of mlo-cgo, mlo-spo and cgo-spo are 0.25, 0.23 and 0.16 ppm respectively. Systematic inter-annual variability is well-defined and is reflected similarly in both IH records, and is consistent with mlo driving most of the variation.

Variations that exceed the ensemble monthly standard deviations include:

- The overall increase in IH difference, generally attributed mainly to increasing NH fossil fuel CO_2 emissions is indicated by a linear regression through the mlo-cgo values (with slope 0.056 ± 0.021 ~~0.061~~ ppm yr^{-1}; mlo-spo gives 0.062 ± 0.021 ~~0.060~~ ppm yr^{-1}). The slope of regressions is much higher for the Apr-May data (0.087 ± 0.011 ~~0.09~~ ppm yr^{-1}) than for Sep-Oct data (0.049 ± 0.011~~0.05~~ ppm yr^{-1})~~, slightly less than the overall mean~~.

- From 1992-2017, the majority of minima occur in Sep; of 26 minima, 24 occur in Sep and 2 in Oct (1992 & 1995). Of the 26 maxima 20 are in May, and 6 in Apr (1997, 1999, 2000, 2004, 2005, 2016).

- Scatter in the amplitude of seasonal maxima (boreal winter/spring) is smaller before 1999. The step-like behaviour in Apr-May from 2009 to 2010 remains the major anomaly.

- In contrast, the minima (in boreal summer/autumn) exhibit greater scatter before 2011, replaced afterwards by a smooth decline to a marked 2016 minimum, then sudden reset in 2017.

- Unusually low boreal summer/autumn IH minima also occur in 1993-1994. Apart from being a period when measurement and calibration methods were consolidating (as discussed next section) the most significant volcanic influence (Pinatubo) is potentially an influence at this time.

A question arises ~~This raises a question~~ as to how well mlo data represents the N~~orthern~~ ~~H~~emisphere. Of more relevance to this study is how well do the mlo samples represent air that is transferred into the Southern Hemisphere. Flasks are collected at mlo above 3 km altitude in down-slope winds, close to the upper troposphere regions where the IH transfer processes defined in FF18 occur (see Figure 5 below), circumstances not shared by other NH surface monitoring sites.

Unlike in typical growth rate analyses, the peak and trough values are largely independent. This is visually explored in Figure 3 using plotting software spline polylines linking peaks (solid) and dips (dashed) months of IH ΔCO_2~~IH CO$_2$ differences~~. Typical trace gas mixing within extra-tropical hemispheres is typically estimated at 1-2 months or less, and inter-hemispheric exchange times estimated at 6-12 months or more (e.g. Bowman and Cowan, 1997, Jacob, 1999). Monthly changes in the peak and trough IH ΔCO_2~~IH CO$_2$~~ largely reflect flux changes in or out of the extra-tropical northern troposphere close to that month. The following sections seek similarities with possible causal forcing processes.

5   Processes influencing CO_2 IH difference variations

Global carbon cycle models generally attribute short term variations in atmospheric CO_2 to exchanges with the terrestrial biosphere (Le Quéré et al., 2018; Rödenbeck et al., 2018; Yue et al., 2017) and implicitly assume model atmospheric transport is correct on all time frames. ~~Variations in atmospheric CO$_2$ global carbon cycle models generally attribute short term variations in the global budget to exchanges with the terrestrial biosphere (Le Quéré et al., 2018; Yue et al., 2017) and implicitly assume model atmospheric transport is correct.~~ While the models have demonstrated an impressive ability to predict mid-to-high latitude CO_2 variations influenced by weather, it is less clear that short term variations in IH exchange (of a magnitude sufficient to influence hemispheric growth rates) have been adequately captured.

~~To exploit the sharp definition of the peaks and dips of IH CO$_2$ in Figure 3, their amplitude and timing are initially compared to anomalies in monthly estimates of terrestrial biosphere emissions and wildfires in Figure 4. It is~~

~~assumed here that flux variations from ocean sources are much smaller than terrestrial and IH atmospheric flux variations and can be generally neglected.~~

5.1 Air-Surface fluxes influencing IH ΔCO$_2$ ~~Terrestrial fluxes influencing IH Transport~~:

The relative magnitude and timing of monthly variations IH ΔCO$_2$ are compared to those in the terrestrial biosphere, wild fires and fossil fuel. (Possible contributions from air-sea exchange are discussed below in relation to Figure 8.)

The primary determinant of the well-defined seasonality in IH ΔCO$_2$ ~~IH CO$_2$~~ in Figure 3 is widely attributed to the temperature-moderated photosynthesis/respiration cycle of NH forests. ~~The influence of equatorial vegetation surface-air CO$_2$ exchange that mixes into both hemispheres is suppressed in IH CO$_2$. The issue here is whether anomalies in the IH CO$_2$ are related to anomalies in terrestrial exchange.~~

Monthly Dynamic Vegetation Model (DVM) estimates of Net Terrestrial Biosphere Production (NBP) in three latitude bands 90ºN-30ºN, 30ºN to 30ºS and 30ºS-90ºS over the 1992-2016 period are obtained with the Community Atmosphere Biosphere Land Exchange (CABLE) model (Kowalczyk et al., 2006; Haverd et al., 2018) ~~are used~~. In addition, extra-tropical (ET) NBP from an ensemble of 16 Land Surface Models (shown in Figure 2 of Bastos et al., 2018) are considered. Because of the small SH contribution, the ensemble ET values are most comparable to CABLE NH NBP. Note: We do not discuss air-surface fluxes derived from CO$_2$ data that are less spatially representative, and/or rely on atmospheric transport modelling. The latter introduce additional model degrees of freedom and potentially overestimate terrestrial variability if the variability in atmospheric IH transport is not adequately captured.

NBP signs are reversed and are described as terrestrial-to-air carbon fluxes. Global wildfire emissions from the Global Fire Emissions Database (Randerson et al., 2018, GFED4.1) from 1997-2015 are classified as NH, EQ and EQ/SH~~,~~. ~~and are also examined.~~ Seasonal anthropogenic emission anomalies are calculated as differences from the detrended 2000 to 2016 monthly data of Oda et al., (2017). For each data set, anomalies (in PgC month^{-1}) in seasonal behaviour for each latitude band were determined by subtracting the mean seasonality from the monthly value.

The major seasonal anomalies in NBP and Wildfire emissions that potentially influence IH ΔCO$_2$ ~~IH CO$_2$~~ are shown in Figures 4(a) and 4(b). The largest anomalous surface to air flux is the extreme equatorial emission anomaly from equatorial wildfire in late 1997 (~0.9 PgC over 3 months) and is not associated with unusual behaviour in the IH ΔCO$_2$ ~~IH CO$_2$~~ records.

Despite mixing of CO$_2$ within the extra tropical Northern Hemisphere being as rapid as 1-2 weeks (Jacob, 1999) compared to IH exchange times of greater than 6 months (Bowman and Cowan, 1997), we see strong correlations with transport for unlagged 3-month averages. And since IH ΔCO$_2$ peaks re-occur within 1 month of the same time each year, close correspondence in timing of terrestrial anomalies and the IH ΔCO$_2$ peaks would be expected if NH terrestrial exchange was the main determinant. This is not evident in Figure 4. More importantly, the amplitude range of terrestrial anomalies appears to be far too small to account for the magnitude of the changes in the peaks and dips of IH ΔCO$_2$.

~~With rapid mixing within the extra tropical northern hemisphere, as rapidly as 1-2 weeks (Jacob, 1999), and since IH CO$_2$ peaks re-occurring at the same time within 1 month of each year, if terrestrial NH emissions were a major~~

~~determinant, close correspondence in timing of terrestrial anomalies and the IH CO_2 peaks would be expected. This is not evident in Figure 4. More importantly, the amplitude of terrestrial anomalies appear to be far too small to influence the IH CO_2 peaks and dips changes.~~

Over the last 25 years the annual relationship between global (mainly NH) fossil fuel combustion emissions and IH difference has been 2.8 PgC ppm^{-1} (0.36 ppm (PgC)$^{-1}$, FF18, see also Figure 9 below). This is when northern fossil fuel emissions effectively mix globally. The volume of the troposphere north of Mauna Loa is around 33% of the global troposphere, so that on the shorter time frame of within-hemisphere mixing, only ~0.92 PgC is required to change the NH background CO_2 by 1 ppm. In Figure 4(c) we round this to 1 PgC = 1 ppm for simplicity.

The variability in the air-surface fluxes, relative to that in IH ΔCO_2, is displayed in Figure 4(d), which plots the standard deviations of residuals from the mean seasonality, for each month over the available record. As for the peaks and dips, we assume a 1:1 relationship between ppm and PgC month^{-1} in IH ΔCO_2. The main variation in IH ΔCO_2 occurs in Mar-Apr, when variability in surface-air fluxes is small but variability in eddy IH exchange is large (see below). A second peak in IH ΔCO_2 standard deviation occurs Aug-Sep, around the time of the dips (but also when equatorial wild fires are more active suggesting a possible contribution from the equatorial emissions at this time?).

~~The IH CO_2 fluxes necessary to explain variations in the peaks (Apr-May) and dips (Sep-Oct) remain almost an order of magnitude greater than the anomalies in estimated terrestrial biosphere and wildfire fluxes.~~

Accepting the precision and near-hemispheric spatial representation of the composite IH ΔCO_2 ~~IH CO_2~~ records, these inconsistences with terrestrial emissions in both timing and magnitude suggest that there are other short-term influences on IH ΔCO_2 ~~IH CO_2~~ of greater magnitude than terrestrial exchange.

5.2 Wind indices reflecting IH Transport:

In contrast to the case for air-surface exchanges, there are a number of prominent features in the composite ~~IH CO_2~~ IH ΔCO_2 records that are shared with behaviour in the dynamical indices of FF18. Inter-hemispheric exchange of CO_2 occurs mainly by eddy processes in the boreal winter-spring and by mean convection and advection associated with the Hadley circulation in the boreal summer-autumn (FF18 and references therein). FF18 developed wind indices that characterize both types of IH transport based on reanalysis data sets focusing on the National Center for Environmental Prediction (NCEP) and National Center for Atmospheric Research (NCAR) reanalysis (NRR) data (Kalnay et al., 1996). Eddy transport is described by u_{duct}, the average 300 hPa zonal velocity in the Pacific Westerly duct region (Frederiksen and Webster, 1988) of -5ºN to 5ºS, 140 to 170ºW (FF16, FF18). Here we use that index and two of the four indices for mean transport introduced in FF18. These are ω_P, the average 300 hPa vertical velocity in pressure coordinates in the region 10-15ºN, 120 to 240ºE, and v_P the average 200 hPa meridional velocity in the region 5-10ºN, 120 to 240ºE. Figure 5 provides a schematic of the geographical location of regions used by FF18, and time series of the monthly values of wind indices are shown in Figure 6.

The top panel in Figure 6a shows a 3-decade time series of the u_{duct} index which characterizes cross-equatorial Rossby wave dispersion, Rossby wave breaking and corresponding increases in transient kinetic energy and eddy transport in the near-equatorial upper troposphere (Webster and Holton, 1982, Frederiksen and Webster, 1988,

Ortega et al., 2018). The large scale Rossby waves are generated by thermal anomalies and topographic features including the Himalayan mountains from which they propagate south-eastward and are able to penetrate into the SH when u_{duct} is positive, corresponding to an open Pacific Westerly duct.

The ω_P and v_P indices in Figures 6b and 6c describe the strength of the mean transport by the Hadley cell in the Pacific region with negative ω_P corresponding to uplift and negative v_P to north to south transport.

Net interhemispheric trace gas exchange requires a partial pressure difference between hemispheres. For CO_2 the average seasonal cycle of 25-year mean partial pressure difference, represented here by monthly baseline mlo-cgo, is shown in Figure 7(a) (mlo-spo is not shown since, reflecting on data quality, it is effectively identical).

The positive mean IH ΔCO_2 ~~IH CO₂~~ is largely due to fossil fuel emissions. Months of positive (north-south) IH difference are shaded green and only in Sep-Oct is there a small reverse gradient. Transport of CO_2 from the northern to the southern hemisphere occurs when green shaded areas in Fig. 7(a) coincide (on average) with blue shaded areas (Fig. 7(b), via eddy transfer with index u_{duct}), or with red shaded areas (Figs. 7(c) and 7(d), via mean transport with indices ω_P and v_P).

Figure 7 also demonstrates that differences from the long-term mean in transport indices (average for each month) vary between the significant El Niño events in 1998, 2010 and 2016:

- In 2010, the IH ΔCO_2 ~~IH CO₂ difference~~ exceeds the average between Feb-Jul (Fig. 7(a)) with reduced eddy transfer between Feb-Apr, associated with lower that average u_{duct} (Fig. 7(b)). Further, between Jun-Sep, there is weaker ascent (Fig. 7(c)) and north to south upper tropospheric wind (Fig. 7(d)) in the key regions defining ω_P and v_P. As noted in FF18, the IH ΔCO_2 ~~IH CO₂~~ eddy and mean transports reinforce to contribute to the unprecedented 2009 to 2010 step in IH ΔCO_2 ~~IH CO₂~~ gradient.

- In 2016, the IH ΔCO_2 ~~IH CO₂ gradient~~ is larger than average between Feb-Jun and smaller than average between Jul-Oct (Fig. 7(a)). These results are again consistent with the behaviour of the dynamical indices. There is reduced IH ΔCO_2 ~~IH CO₂~~ eddy transfer in the first half of the year (Fig. 7(b)) but very strong mean transport in the second half of the year (Figs. 7(c) and(d)) that accounts for the annual IH ΔCO_2 ~~IH CO₂ gradient~~, as noted in FF18.

- In 1998, the IH ΔCO_2 ~~IH CO₂ difference~~ exceeds the average from May-Dec and is close to the mean annual cycle for the rest of the year. We note from Figure 7 that the annual increase in IH ΔCO_2 ~~IH CO₂ gradient,~~ also shown in Figure 2 of FF18, is largely induced by the Jun-Aug mean Hadley circulation.

- It is suggestive that the relative variation in IH ΔCO_2 ~~IH CO₂~~ Feb-May for the three big El Niño years matches that in u_{duct}, however it is puzzling that the largest u_{duct} anomaly, 1998, is when IH ΔCO_2 ~~IH CO₂~~ is closest to the mean behaviour. The fact that the mean transport indices at this time of year are also consistently well below their long term average ~~is also~~ is also of note, since ~~with u_{duct}~~ with u_{duct} close to zero and ~~-ω-~~ $-\omega_P$, $-v_P$ indicating descent and south to north meridional winds, there is no obvious mechanism for IH exchange in this season. Yet, over the 25 years correlation of the Apr-May IH ΔCO_2 ~~IH CO₂~~ peaks with ~~-ω-~~ $-\omega_P$, $-v_P$ is significant, r ≈ 0.4. One possible explanation for these behaviours in the early part of the Boreal winter/Austral summer may be found in changes in the volume of the well-mixed portion of the northern hemisphere (see Discussion, Section ~~5~~7).

Different responses of IH ΔCO_2 ~~IH CO₂~~ to wind indices at different ENSO events, and from non-ENSO periods, are discussed in Section ~~6~~7.

As an aside, we also include a similar plot for the average SH cgo-spo differences in Figure 8. Despite some concerns about flask storage at spo (e.g. due to long flask-air storage times), all networks indicate that on average spo baseline CO_2 exceeds that at cgo in the austral summer months. The minimum cgo-spo appears to precede inversion estimates of Southern Ocean CO_2 uptake south of 30ºS (Lenton et al., 2013). High precision continuous CO_2 monitoring across the Southern Ocean (Stavert et al., 2019, and personal communication) confirm small and relatively smooth seasonal variation. The earlier Nov-Dec minimum in CO_2 difference coincides with a seasonal dip in fossil fuel emissions (Oda et al., 2017) perhaps indicating an alternative explanation.~~It suggests a possible role for southern ocean uptake influencing cgo, although the Oct-Feb maximum influence (minimum cgo-spo and sea-air flux) in Nov-Dec appears to precede inversion estimates of Southern Ocean CO_2 uptake south of 30ºS. (Lenton et al., 2013).~~

6   Year-to-year variation in the composite records

The annual net impacts of the various potential influences on site IH ΔCO_2 ~~IH CO_2 differences~~ (when typical terrestrial biosphere seasonal variations are balanced) appear in Figure 9. The typical uncertainty in annual values obtained from combining composite standard deviations of normalised monthly values is ±0.03 ppm, compared to ~±0.3 ppm variation in the detrended annual record.

Working through Figure 9 from the left in order to highlight other systematic features:

~~The largest cgo-spo differences occur in 1992 and 1993. In these years the composite cgo-spo involves uninterrupted monthly data from two networks only, NOAA and CSIRO. At this time, CSIRO co-measurement of a subset of NOAA cgo flask samples (Masarie et al., 2001) show CSIRO cgo is around +0.2 ppm higher than NOAA cgo, too small and in the wrong direction to explain the 1992-1993 network difference, rather favouring a temporary problem with NOAA spo sampling; no such persisting anomalies of this magnitude have occurred since. Prior to Mar 1993, NOAA cgo-spo data are lower by around -0.5 ppm compared to CSIRO data (Supplement Figure S4), with much larger seasonality compared to subsequent NOAA & CSIRO data.~~

~~Working through Figure 9 from the left in order to highlight other systematic features:~~

- Except for 2016, every major El Niño (as indicated by the magnitude of the peak Oceanic Niño Index, for ONI > 1) corresponds to a transition from low to high IH difference (conclusions are similar using Nino3 or Nino3.4).~~;~~ ~~H~~however, the CO_2 response is not proportional to ONI, e.g. comparing 2009-10 to 1997-98, or most noticeably to 2015-16 (the strongest ONI but the smallest IH ΔCO_2 ~~IH CO_2~~ step).

- At the strong La Niña events (ONI<-1) there is no clear indication of anomalous IH ΔCO_2.

- There is remarkable stability in IH ΔCO_2 ~~IH CO_2~~ from 2010-2014 (despite the strong La Niña in 2011).~~;~~ ~~after~~ After 2010, there are no significantly positive ONI anomalies (El Niños), and the 5-year increase of ~0.1 ppm is lower than that generally attributed to the increasing ~~upward~~ mean fossil fuel emissions (~~(~~the 2010-2014 change in FF is 0.73 PgC⁻¹, which at 0.36 ppm~~(~~ (PgC)⁻¹ would result in a 0.26 ppm increase)~~)~~. The FF with this scaling is shown at the top. There is markedly less variability (the composite standard deviation of de-trended annual means ~is 0.04 ppm) than any equivalent period over the previous 16 years (~0.31 ppm).

- The 2009/2010 year-to-year change of ~0.8 ppm (addressed in FF16 using CSIRO data only) remains the major year-to-year change in the annual records. The current composite data confirm the general FF16 conclusion. ~~Subsequent data now offer an improved perspective on FF16 conclusions.~~

-

- The linear regression through the 25 year mlo-cgo annual data gives a slope of 0.067 ± 0.006 ~~0.72~~ ppm yr^{-1} compared to that through monthly values of 0.56 ± 0.021 ~~0.61~~ ppm yr^{-1} in Figure 3, or through the peaks of 0.087 ± 0.011 ~~0.09~~ ppm yr^{-1} or the dips of 0.049 ± 0.011 ~~0.05~~ ppm yr^{-1}. We interpret this as indicating the combined long-term influence of both eddy and mean transport on the annual mean IH ΔCO_2 ~~IH CO₂~~.

- In 2017, the IH difference is close to the long term mean, with the duct open and Hadley strength returning to be close to its long-term mean

- ~~In 2017, the IH difference is close to the long term mean, with the duct open and Hadley strength returning to be close to its long-term mean.~~

-

7 Discussion

The composite monthly IH differences reveal variation from monthly to decadal time scales that exceed measurement and sampling error (as indicated by the composite standard deviations) thus requiring biogeochemical explanation. This discussion focusses on the potential of IH transport measured by wind indices to explain major features in IH ΔCO_2 variation, with emphasis on periods and events when they are likely to be the dominant influence on IH ΔCO_2 ~~IH CO₂~~. It complements ~~is intended as a complement to~~ the more general statistical analyses in FF18.

In Figure 6, decreasing u_{duct} acts to lessen eddy IH exchange and increase IH ΔCO_2 ~~IH CO₂~~, while the increasing Hadley circulation (decreasing v_P and ω_P) decreases IH ΔCO_2 ~~IH CO₂~~.

The fact that the magnitude of IH ΔCO_2 ~~IH CO₂~~ response varies greatly between the 1998, 2010 and 2016 El Niño events (with little or no eddy transfer occurring in boreal winter/spring in these years) is consistent with a quasi-decadal variation in the negative excursions of v_P and ω_P in Figure 6 (most obvious in ω_P). In 1998 and 2010, the Hadley boreal summer/autumn indices are closer to zero, while 2016 registers an unprecedented negative excursion.

The complication of IH ΔCO_2 ~~IH CO₂~~ variations in the boreal winter/austral summer when u_{duct}, ω_p and v_P indices indicate that little or no IH exchange occurs (and u_{duct} closure tends to increase mlo CO_2), is at a time when the north to south seasonal variation in the Inter-Tropical-Convergence-Zone is near maximum. If NH peak terrestrial emissions (biospheric and industrial) at that time are diluted into a larger volume of well-mixed NH air, it could offset the mlo CO_2 increase anticipated from u_{duct} closure. This volume effect is likely to be a second order effect in non-El Niño years.

NH terrestrial biosphere emission anomalies in the 2010-2014 period (Figure 4) are more variable than those in 2000-2005, the opposite of the relative behaviour in IH ΔCO_2 ~~IH CO₂~~ variability in Figure 9. These emissions are relatively small, and frequently occur after the larger IH ΔCO_2 ~~IH CO₂~~ anomalies, all inconsistent with a significant contribution to the composite IH differences; so they are considered second order. The small 2010-

2014 trend (~0.1 ppm compared to 0.26 ppm expected from fossil fuel emissions), and the steadily decreasing westerly wind strength in u_{duct} over the period, should increase IH ΔCO_2 ~~IH CO₂~~ over the fossil fuel trend (FF16). The flattening trend is consistent with the IH ΔCO_2 ~~IH CO₂~~ flux due to IH mixing by the Hadley process overwhelming the increases expected from Fossil Fuel combustion and from decreasing u_{duct} strength. There is a linear relationship between u_{duct} and equatorial upper troposphere transient kinetic energy shown in Fig. 6 of Frederiksen and Webster 1988 and discussed in FF16 and FF18). Note that in Figure 6, there is no precedent for similar sustained opposing behaviour in the two modes of IH transfer. The trend and lack of scatter in 2010-2014 IH ΔCO_2 ~~IH CO₂~~ can be understood by the IH ΔCO_2 ~~IH CO₂~~ fluxes being significantly larger than air-surface exchanges at the time.

The magnitude of the IH flux anomalies of up to ~2 PgC month^{-1} exceed known air-surface fluxes in the NH and are of a magnitude to significantly influence NH CO_2 growth rate variability. With increasing fossil fuel fluxes, the role of IH exchange on the IH ΔCO_2 ~~CO₂ IH differences~~, and NH CO_2 growth, is expected to become increasingly important.

The previous inability of carbon cycle models to simulate the 2009/2010 step (FF16, Supplementary Information) implies that there is inadequate parameterisation of IH ΔCO_2 ~~IH CO₂~~ transfer, particularly by eddy exchange, in some global carbon cycle models. If this is the case, then studies that interpret CO_2 behaviour during ENSO events as a guide to terrestrial biosphere responses to climate (e.g. Rödenbeck et al., 2018) will also be compromised. The ability to simulate the identified features of the composite IH ΔCO_2 ~~IH CO₂ differences~~ (within the standard deviations) would provide convincing independent confirmation of atmospheric transport implementation.

7   Conclusions

Over the last 25-years there is a high degree of agreement in the measurement of monthly spatial differences in background CO_2 levels by three measurement laboratories using four different sampling methodologies and sampling frequencies. Geographic isolation of sample collection sites and consistent sophisticated background selection over the 25 years, as well as coincident monitoring of a wide range of atmospheric species, excludes local and regional influence on CO_2 at mlo, spo and cgo to an extent not generally available at other surface monitoring sites.

The temporal variation in the composite IH ΔCO_2 ~~IH CO₂ differences~~ exhibit several systematic features on monthly to multi-year timeframes that are not reflected in independent evidence of air-surface exchange, but do correspond to features in dynamical indices selected to represent both eddy and mean IH exchange. The comparisons in this paper imply a major role for IH exchange of CO_2 in NH growth rate variations.

The evidence for a significant influence of atmospheric dynamics on the IH ΔCO_2 ~~CO₂ IH~~ gradient has relevance for global carbon cycle studies. It implies that both eddy and mean transport processes, and volume effects, need to be specifically included in transport model simulations, since the balance between the two is constantly changing, particularly in El Niño periods when eddy transport is reduced. It also means that El Niño events may be a poor predictor of the carbon cycle behaviour in non-ENSO years.

Global carbon cycle model simulations should be able to reproduce the major features identified here in the composite IH records if the re-analyses transport is correctly implemented. In attempting to simulate the composite

differences, one complication is model selection of baseline that matches the flask sampling criteria. While monthly baseline averages appear to succeed in this respect, a more comprehensive treatment (outside the scope of this study) based on individual flask measurements rather than monthly averages, and other trace gas observations FF16, FF18, and in particular radon (Chambers et al., 2016), could possibly improve this process.

The Supplement related to this article is provided

Author contributions.

RJF generated the composite records and their analyses while JSF provided information on atmospheric dynamics and the roles of transport mechanisms. LPS and RLL contributed CO_2 measurement quality assessments. All four authors contributed to the written document.

Competing interests. The authors declare that they have no conflicts of interest.

References

Bastos, A., Freidlingstein, P., Sitch, S., Chi Chen, Mialon, A. Wigneron, J-P., Arora, V. K.,, Briggs, P.R., Canadell, J. G., Ciais, P., Chevallie, F., Lei Cheng, Delire, C., Haverd, V., Jain, A. K., Joos, F., Kato, E., Lienert, S., Lombardozzi, D., Melto, J. R.,Myneni, R., Nabel, J. E. M. S., Pongratz, J., Poulter, B., Rödenbeck, C., Séférian, R., Tian, H., van Eck, C., Viovy, N., Vuichard, N., Walker, A. P., Wiltshire, A., Jia Yang, Zaehle, S., Ning Zeng, and Zhu, D: Impact of the 2015/2016 El Niño on the terrestrial carbon cycle constrained by bottom-up and top-down approaches. Phil. Trans. R. Soc. B 373: 20170304. http://dx.doi.org/rstb.2017.0304, 2018.

Bowman K. P. and Cohen, P. J.: Interhemispheric exchange by seasonal modulation of the Hadley Circulation, J. Atmos. Sci. 54, 2045-2059, 1997.

Chambers, S. D., Williams, A. G., Conen , F., Griffiths, A. D., Reimann, S. , Steinbacher , M., Krummel, P. B., Steele, L. P., van der Schoot, M. V., Galbally, I. E., Molloy ,S. B. and Barnes J. E.: Towards a Universal "Baseline" Characterisation of Air Masses for High- and Low-Altitude Observing Stations Using Radon-222, Aerosol and Air Quality Research, 16: 885–899, 2016, ISSN: 1680-8584 print / 2071-1409 online doi: 10.4209/aaqr.2015.06.0391, 2016.

Conway, T. J., Tans, P. P., Waterman, L. S., Thoning, K. W., Kitzis, D. R., Masarie, K. A. and Ni Zhang: Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, J. Geophys. Res., 99, 22,831-22,855, 1994.

Dlugokencky, E. J., Lang, P.M., Masarie, K.A., Crotwell, A.M., and Crotwell, M.J.: Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1968-2013, Version: 2014-06-27, 2014. Path: ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/surface/

Dunse, B. L., Steele, L. P., Fraser, P. J. and Wilson, S. R.: An analysis of Melbourne pollution episodes observed at Cape Grim from 1995-1998, in *Baseline Atmospheric Program (Australia) 1997-1998*, edited by N. W. Tindale, N. Derek, and R. J. Francey, Bureau of Meteorology and CSIRO Atmospheric Research, Melbourne, Australia, 34-42, 2001.

Francey, R.J., Steele, L. P., Langenfelds, R. L., Lucarelli, M., Allison, C. E., Beardsmore, D. J. Coram, S. A., Derek, N., de Silva, F. R., Etheridge, D. M., Fraser, P. J. , Henry, R. J., Turner, B. and Welch, E. D.: Global Atmospheric Sampling Laboratory (GASLAB): supporting and extending the Cape Grim trace gas programs. *Baseline Atmospheric Program (Australia) 1993*. (eds. R.J. Francey, A.L.Dick and N.Derek) Bureau of Meteorology and CSIRO Division of Atmospheric Research, Melbourne, pp 8-29 (1996).

Francey, R. J., & Frederiksen, J. S.: The 2009–2010 step in atmospheric CO_2 interhemispheric difference, Biogeosciences*, 13, 873- 885, doi: 10.5194/bg-13-873-2016, 2016.

Frederiksen, J. S. & Webster, P. J.: Alternative theories of atmospheric teleconnections and low-frequency fluctuations, Rev. Geophys., 26, 459-494, 1988.

Frederiksen, J. S. and Francey, R. J.: Unprecedented strength of Hadley circulation in 2015–2016 impacts on CO_2 interhemispheric difference, Atmos. Chem. Phys., 18, 14837-14850, https://doi.org/10.5194/acp-18-14837-2018, 2018.

Haverd, V., Smith, B., Nieradzik, L., Briggs, P. R., Woodgate, W., Trudinger, C. M., Canadell, J. G., and Cuntz, M.: A new version of the CABLE land surface model (Subversion revision r4601) incorporating land use and land cover change, woody vegetation demography, and a novel optimisation-based approach to plant coordination of photosynthesis, Geosci. Model Dev., 11, 2995-3026, https://doi.org/10.5194/gmd-11-2995-2018.

Jacob, D. J.: Introduction to Atmospheric Chemistry, Princeton University Press, 1999.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., & Joseph, D.: The NCEP/NCAR Reanalysis 40-year Project, Bull. Amer. Meteor. Soc., 77, 437-471, 1996

Keeling, R.F. and Schertz, S.R.: Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle, Nature, 358, 723-727, 1992.

Keeling C.D.: Rewards and penalties of monitoring the earth, *Annu. Rev. Energy Environ.* 1998. 23:25–82, 1998.

Keeling, C. D., Piper, S. C., Bacastow, R. B., Wahlen, M., Whorf, T. P., Heimann, M. and Meijer, H. A.: Exchanges of atmospheric CO_2 and $^{13}CO_2$ with the terrestrial biosphere and oceans from 1978 to 2000. I. Global aspects, SIO Reference Series, No. 01-06, Scripps Institution of Oceanography, San Diego, 88 pages, 2001.

Kowalczyk, E. A., Wang Y. P., Law, R. M., Davies, H. L. McGregor, J. L. and Abramowitz, G.: The CSIRO Atmosphere Biosphere Land Exchange (CABLE) model for use in climate models and as an offline model. National Library of Australia Cataloguing-in-Publication, ISBN 1 921232 39 0 (pdf.), 2006.

Langenfelds, R. L., Francey, R. J., Steele, L. P., Dunse, B. L., Butler, T. M., Spencer, D. A., Kivlinghon, L. M., and Meyer, C. P.: Flask sampling from Cape Grim overflights, in *Baseline Atmospheric Program (Australia) 1999-2000*, edited by N. W. Tindale, N. Derek and P. J. Fraser, Bureau of Meteorology and CSIRO Atmospheric Research, Melbourne, Australia, 73-75, 2003.

Lenton, A., Tilbrook, B., Law, R. M., Bakker, D., Doney, S. C., Gruber, N., Ishii, M., Hoppema, M., Lovenduski, N. S., Matear, R. J., McNeil, B. I., Metzl, N., Mikaloff Fletcher, S. E., Monteiro, P. M. S., Rödenbeck, C., Sweeney, C., and Takahashi, T.: Sea–air CO_2 fluxes in the Southern Ocean for the period 1990–2009, Biogeosciences, 10, 4037-4054, https://doi.org/10.5194/bg-10-4037-2013, 2013.

Lintner, B. R., Gilliland, A. B. and Fung, I. Y.: Mechanisms of convection-induced modulation of passive tracer interhemispheric transport interannual variability, J. Geophys. Res., 109, D13102, doi:10.1029/2003JD004306, 2004

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., Korsbakken, J. I., Peters, G. P., Canadell, J. G., Jackson, R. B., Boden, T. A., Tans, P. P., Andrews, O. D., Arora, V. K., Bakker, D. C. E., Barbero, L., Becker, M., Betts, R. A., Bopp, L., Chevallier, F., Chini, L. P., Ciais, P., Cosca, C. E., Cross, J., Currie, K., Gasser, T., Harris, I., Hauck, J., Haverd, V., Houghton, R. A., Hunt, C. W., Hurtt, G., Ilyina, T., Jain, A. K., Kato, E., Kautz, M., Keeling, R. F., Klein Goldewijk, K., Körtzinger, A., Landschützer, P., Lefèvre, N., Lenton, A., Lienert, S., Lima, I., Lombardozzi, D., Metzl, N., Millero, F., Monteiro, P. M. S., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Nojiri, Y., Padin, X. A., Peregon, A., Pfeil, B., Pierrot, D., Poulter, B., Rehder, G., Reimer, J., Rödenbeck, C., Schwinger, J., Séférian, R., Skjelvan, I., Stocker, B. D.,

Tian, H., Tilbrook, B., Tubiello, F. N., van der Laan-Luijkx, I. T., van der Werf, G. R., van Heuven, S., Viovy, N., Vuichard, N., Walker, A. P., Watson, A. J., Wiltshire, A. J., Zaehle, S., and Zhu, D.: Global Carbon Budget 2017, Earth Syst. Sci. Data, 10, 405–448, https://doi.org/10.5194/essd-10-405-2018, 2018.

Masarie, K. A., Langenfelds, R.L., Allison, C.E., Conway, T.J., Dlugokencky, E.J., Francey, R.J., Novelli, P.C., Steele, L.P., Tans, P.P., Vaughn, B. and White, J.W.C.: NOAA/CSIRO Flask Air Intercomparison Experiment: A strategy for directly assessing consistency among atmospheric measurements made by independent laboratories, Journal of Geophysical Research, 106 (D17): 20445-20464, 2001.

Oda, T., Maksyutov, S., and Andres, R. J.: The Open-source Data Inventory for Anthropogenic Carbon dioxide (CO_2), version 2016 (ODIAC2016): A global, monthly fossil-fuel CO2 gridded emission data product for tracer transport simulations and surface flux inversions, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2017-76, in review, 2017.

Ortega, S., Webster, P. J., Toma, V., & Chang, H. R.: The effect of potential vorticity fluxes on the circulation of the tropical upper troposphere, Quart. J. Roy. Met. Soc., doi: 10.1002/qj.3261, 2018.

Pak, B. C., Langenfelds, R. L., Francey, R. J., Steele, L. P. and Simmonds, I.: A climatology of trace gases from the Cape Grim Overflights, 1992-1995, in *Baseline Atmospheric Program (Australia), 1994-95*, edited by R. J. Francey, A. L. Dick, and N. Derek, Bureau of Meteorology and CSIRO Division of Atmospheric Research, Melbourne, Australia, 41-52, 1996.

Randerson, J.T., van der Werf, G.R., Giglio, L., Collatz, G.J., and Kasibhatla, P.S.: Global Fire Emissions Database, Version 4.1 (GFEDv4). ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/1293, 2018.

Rödenbeck, C., Zaehle, S., Keeling, R., and Heimann, M.: How does the terrestrial carbon exchange respond to inter-annual climatic variations? A quantification based on atmospheric CO_2 data, Biogeosciences, 15, 2481-2498, https://doi.org/10.5194/bg-15-2481-2018, 2018.

Stavert, A. R., Law, R. M., van der Schoot, M., Langenfelds, R. L., Spencer, D. A., Krummel, P. B., Chambers, S. D., Williams, A. G., Werczynski, S., Francey, R. J., and Howden, R. T.: The Macquarie Island [LoFlo2G] high-precision continuous atmospheric carbon dioxide record, Atmos. Meas. Tech., 12, 1103-1121, https://doi.org/10.5194/amt-12-1103-2019, 2019.

Sturm, P., Leuenberger, M., Sirignano, C. R., Neubert, R. E. M., Meijer, H. A. J., Langenfelds, R., Brand, W. A. and Tohjima, Y.: Permeation of atmospheric gases through polymer O-rings used in flasks for air sampling, J. Geophys. Res. 109, D04309, doi:10.1029/2003JD004073, 2004.

Tans, P. P., Conway, T. J., Dlugokencky, E. J., Thoning, K. W., Lang, P. M., Masarie, K. A., Novelli, P., and Waterman, L. S.: 2. Carbon Cycle Division, 2.1 Continuing Programs. Climate Monitoring and Diagnostics Laboratory, 20, 15-24, 1992.

Thoning, K. W., Tans, P P. and Komhyr, W. D.: Atmospheric Carbon Dioxide at Mauna Loa Observatory 2. Analysis of the NOAA GMCC Data, 1974-1985, J. Geophysical Research, 94, D6, 8549-8565, 1989

Webster, P. J., and Holton, J. R.: Cross-equatorial response to mid-latitude forcing in a zonally varying basic state, J. Atmos. Sci., 39, 722-733, 1982.

Yue, C., Ciais, P., Bastos, A., Chevallier, F., Yin, Y., Rödenbeck, C., and Park, T.: Vegetation greenness and land carbon-flux anomalies associated with climate variations: a focus on the year 2015, Atmos. Chem. Phys., 17, 13903-13919, https://doi.org/10.5194/acp-17-13903-2017, 2017

Zhao, C. L. and Tans, P. P.: Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, Journal of Geophysical Research: Atmospheres, 111(D8), 10.1029/2005JD006003, http://dx.doi.org/10.1029/2005JD006003, d08S09, 2006.

Figure Captions

Figure 1: Monthly CO_2 differences (in ppm) from NOAA for sites mlo, cgo and spo and networks SIO1, SIO2 and CSIRO. 5-month running means are highlighted. 1996-2016 mean and standard deviation (bracketed) values are included.

Figure 2: Three-month averaged CO_2 differences between sites for each of the four sampling networks. (Each 3-month average is plotted on Jan 1 of the appropriate year.)

Figures 3: Composite station difference data showing the network ensemble average and standard deviation of monthly CO_2 for (a) mlo-cgo, (b) mlo-spo and (c) cgo-spo (on a doubly expanded scale). Linear regressions through the IH records are black-dotted lines. Spline polylines link peaks (blue, solid) and dips (red, dotted) of the seasonal IH differences. Shaded blue rounded rectangles indicate El Niño periods with strongly anomalous equatorial zonal winds.

Figure 4: Comparison of the timing and amplitude of terrestrial emission anomalies (i.e. mean seasonality subtracted) with variations of the peaks and dips in Figure 3. (a) shows seasonal anomalies in CABLE emissions (dark green) and in 16-DGVM TRENDY extra-tropical emissions (Bastos et al., 2018; light green) and (b) shows GFED4.1 wild fire seasonal anomalies, for NH (green), EQ (pink) and SH (blue, SH/EQ for GFED4.1). In (c) The largest anomalies (CABLE NH, CABLE EQ, and GFED4.1 EQ) on the left xis are compared to the ppm variation in peaks (red) and dips (blue) on separate right axes. The axes scaling equates 1 PgC with 1 ppm (see text). To highlight seasonal differences, Figure 4(d) shows the standard deviation in the seasonal anomalies for each month, including those in anthropogenic emissions (Oda et al., 2017).Comparison of the timing and amplitude of terrestrial emissions with variations of the peaks and dips in Figure 3. (a) seasonal anomalies in CABLE emissions and (b) GFED4.1 wild fire seasonal anomalies, for NH (green), EQ (pink) and SH (blue, SH/EQ for GFED4.1). In (c) The largest anomalies (CABLE NH, CABLE EQ, and GFED4.1 EQ) on the left axis are compared to the ppm variation in peaks (red) and dips (blue) on separate right axes. The axes scaling equates 1 PgC with 1 ppm (see text).

Figure 5: Schematic of the boundaries and altitudes of regions used in FF18 to define wind indices that describe eddy IH transfer (u_{duct}, westerlies positive) and mean transfer (uplift, negative ω_P) and north to south transfer (negative v_P). The shaded area brackets the austral summer extent of the Inter Tropical Convergence Zone ITCZ in the south (blue dash) and boreal summer extent in the north (red dash).

Figure 6: Monthly values of (a) u_{duct}, (b) ω_P and (c) v_P. Shading The shading to zero indicates months of enhanced transport which acts to reduce the IH ΔCO_2IH CO_2 difference. Anomalous dynamical periods are highlighted with gray shaded rectangles.

Figure 7: The monthly averages of dynamical factors governing CO_2 IH Exchange over the last 25 years. (a) detrended CO_2 partial pressure differences mlo-cgo (green), (b) Pacific Eddy transport index u_{duct} (dark blue), (c) Pacific Hadley transport indicated by uplift at 10-15ºN (-ω_P, light red) and (d) North to South transport (-v_P, dark red). On average, coincidence of shading in wind indices and shaded months of IH ΔCO_2 IH CO_2 is a precondition for increased IH mixing (reduced IH gradient). The more anomalous transport years, 1998 (dots), 2010 (dashes) and 2016 (black line) are shown for each wind index, and for mlo-cgo IH ΔCO_2 IH CO_2.

Figure 8: Composite 25-year average of monthly baseline cgo-spo CO_2 (dark blue). Individual network values are shown in orange (NOAA), dark blue (SIO2) and CSIRO (black). Estimates of sea-air CO_2 flux seasonality are shown in light blue.

Figure 9: Annual changes in the baseline CO_2 difference between sites. Interhemispheric differences are plotted on the left axis. The peak magnitudes of strong El Niños (brown, ONI index > 1) and strong La Niñas (purple, ONI index <-1) are indicated. The cgo-spo annual differences are plotted on a doubled right-hand scale. Annual Fossil Fuel emissions from FF18, are shown on the top right axis.Annual changes in the baseline CO_2 difference between sites. Interhemispheric differences are plotted on the left axis. The peak magnitudes of the strong El Niños (3-month Nino Region 3.4 average or ONI index) are indicated. The cgo-spo annual differences are plotted on a doubled right-hand scale. Dashed lines represent linear regressions through the annual average data.

Table

# networks	four	three	two	one
mlo-cgo	-	277	23	-
mlo-spo	268	44	2	-
cgo-spo		257	42	1

Table 1: Number of months of data available for composite differences at the baseline sites.

Figures

Figure 1: Monthly CO_2 differences (in ppm) from NOAA for sites mlo, cgo and spo and networks SIO1, SIO2 and CSIRO. 5-month running means are highlighted. 1996-2016 mean and standard deviation (bracketed) values are included.

Figure 2: Three-month averaged CO_2 differences between sites for each of the four sampling networks. (Each 3-month average is plotted on Jan 1 of the appropriate year.)
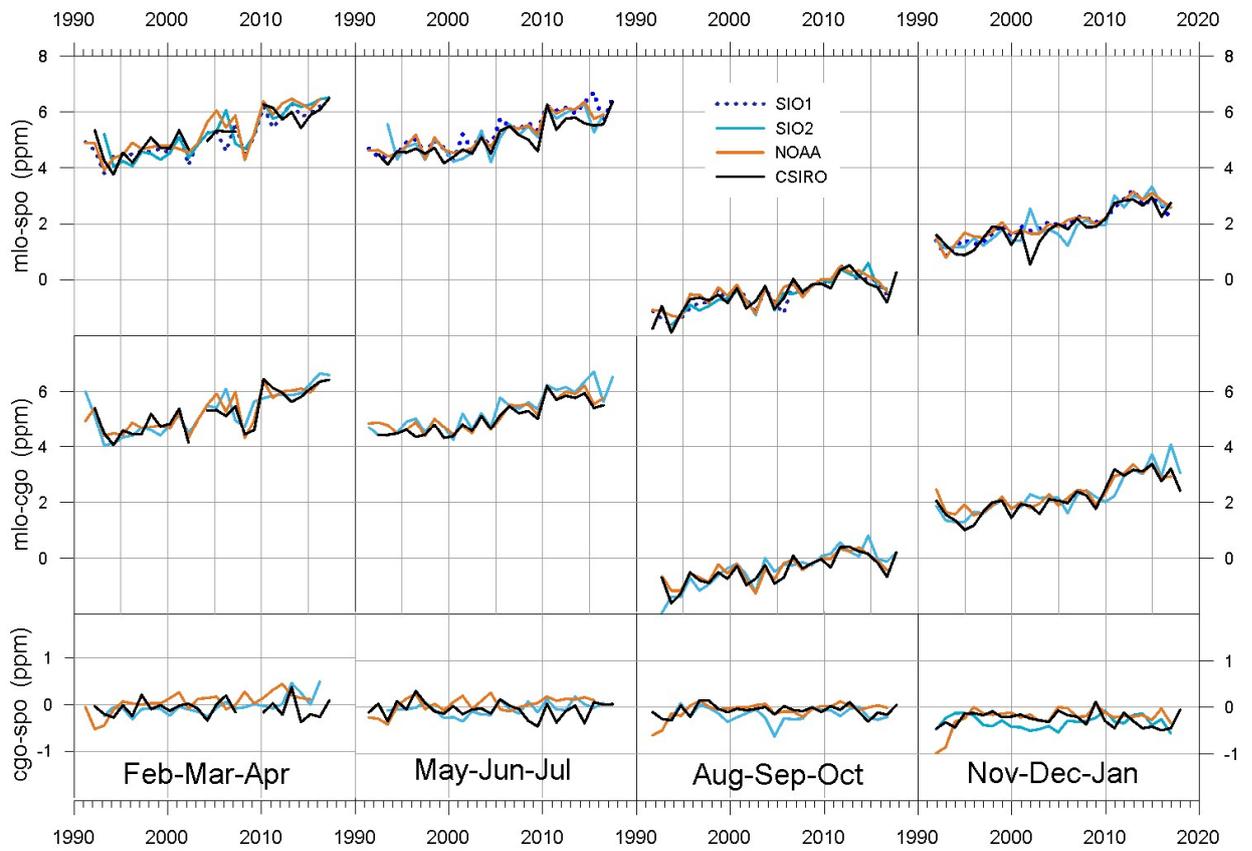
Figures 3: Composite station difference data showing the network ensemble average and standard deviation of monthly CO_2 for (a) mlo-cgo, (b) mlo-spo and (c) cgo-spo (on a doubly expanded scale). Linear regressions through the IH records are black-dotted lines. Spline polylines visually link peaks (blue, solid) and dips (red, dotted) of the seasonal IH differences. Shaded blue rounded rectangles indicate El Niño periods with strongly anomalous equatorial zonal winds.

Figure 4: Comparison of the timing and amplitude of terrestrial emission anomalies (i.e. mean seasonality subtracted) with variations of the peaks and dips in Figure 3. (a) shows seasonal anomalies in CABLE emissions (dark green) and in 16-DGVM TRENDY extra-tropical emissions (Bastos et al., 2018; light green) and (b) shows GFED4.1 wild fire seasonal anomalies, for NH (green), EQ (pink) and SH (blue, SH/EQ for GFED4.1). In (c) The largest anomalies (CABLE NH, CABLE EQ, and GFED4.1 EQ) on the left axis are compared to the ppm variation in peaks (red) and dips (blue) on separate right axes. The axes scaling equates 1 PgC with 1 ppm (see text). To highlight seasonal differences, ~~Comparison of the timing and amplitude of terrestrial emissions with variations of the peaks and dips in Figure 3. (a) seasonal anomalies in CABLE emissions and (b) GFED4.1 wild fire seasonal anomalies, for NH (green), EQ (pink) and SH (blue, SH/EQ for GFED4.1). In (c) The largest anomalies (CABLE NH, CABLE EQ, and GFED4.1 EQ) on the left axis are compared to the ppm variation in peaks (red) and dips (blue) on separate right axes. The axes scaling equates 1 PgC with 1 ppm (see text).~~

Figure 4(d) shows the standard deviation in the seasonal anomalies for each month, including those in anthropogenic emissions (Oda et al., 2017)
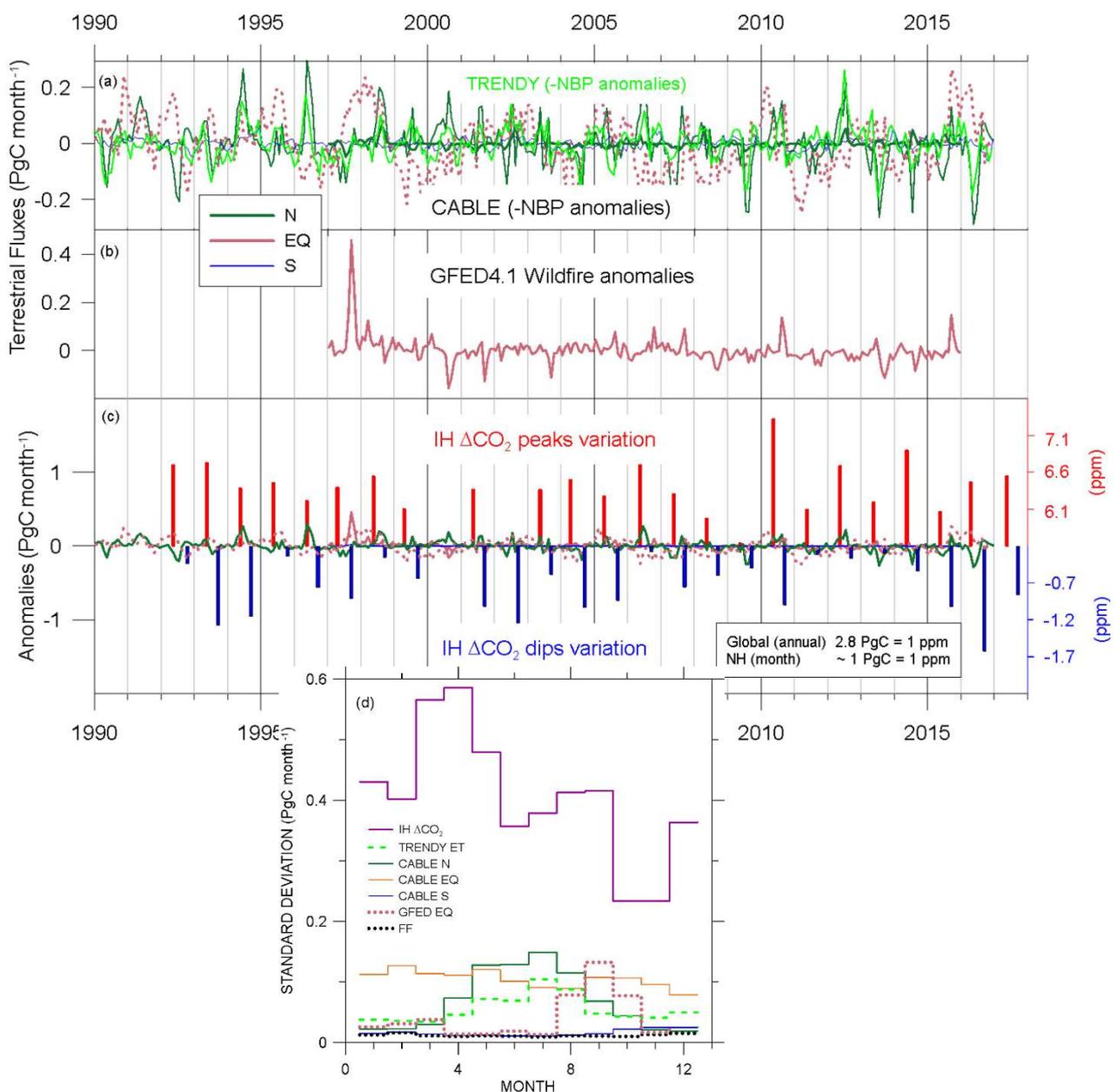
Figure 5: Schematic of the boundaries and altitudes of regions used in FF18 to define wind indices that describe eddy IH transfer ($u_{duct,}$, westerlies positive) and mean transfer (uplift, negative ω_P) and north to south transfer (negative v_P). The shaded area brackets the austral summer extent of the ITCZ in the south (blue dash) and boreal summer extent in the north (red dash).
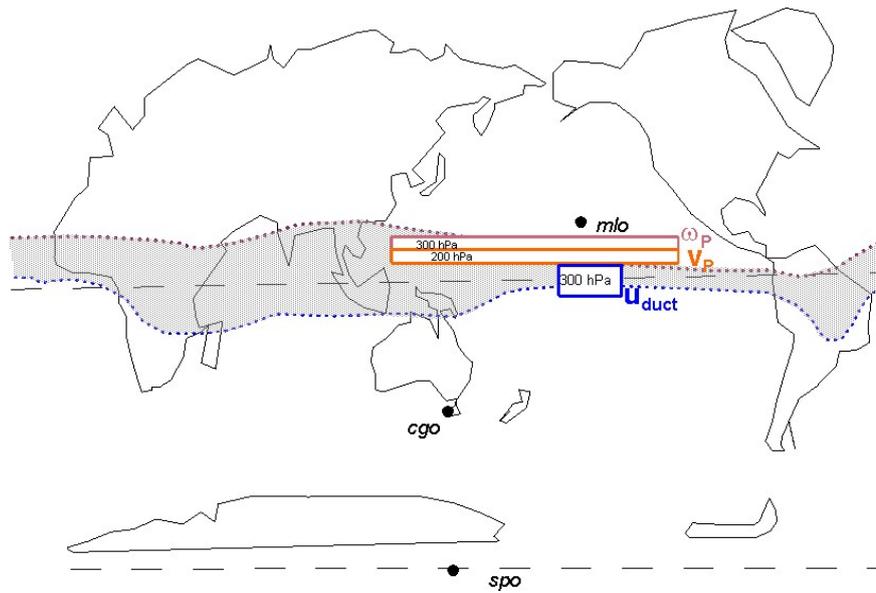
Figure 6: Monthly values of (a) u_{duct}, (b) ω_P and (c) v_P. Shading indicates months of enhanced transport which acts to reduce the ~~IH ΔCO_2~~ ~~IH CO₂ difference~~. Anomalous dynamical periods are highlighted with shaded rectangles.

Figure 7: The monthly averages of dynamical factors governing CO_2 IH Exchange over the last 25 years. (a) detrended CO_2 partial pressure differences mlo-cgo (green), (b) Pacific Eddy transport index u_{duct} (dark blue), (c) Pacific Hadley transport indicated by uplift at 10-15ºN ($-\omega_P$, light red) and (d) North to South transport ($-v_P$, dark red). On average, coincidence of shading in wind indices and shaded months of IH ΔCO_2 ~~IH CO_2~~ is a precondition for increased IH mixing (reduced IH gradient). The more anomalous transport years, 1998 (dots), 2010 (dashes) and 2016 (black line) are shown for each wind index, and for mlo-cgo IH ΔCO_2 ~~IH CO_2~~.

Figure 8: Individual network values of monthly cgo-spo ΔCO_2 that contribute to the composite are shown in orange (NOAA), dark blue (SIO2) and CSIRO (black). Estimates of sea-air CO_2 flux seasonality south of 30°S are shown in light blue. ~~Individual network values that contribute to the composite are shown in orange (NOAA), dark blue (SIO2) and CSIRO (black). Estimates of sea-air CO_2 flux seasonality are shown in light blue.~~

Figure 9: Annual changes in the baseline CO$_2$ difference between sites. Interhemispheric differences are plotted on the left axis. The peak magnitudes of strong El Niños (brown, ONI index > 1) and strong La Ninas (purple,

ONI index <-1) are indicated. The cgo-spo annual differences are plotted on a doubled right-hand scale. Annual Fossil Fuel emissions from FF18, are shown on the top right axis. ~~Annual changes in the baseline CO_2 difference between sites. Interhemispheric differences are plotted on the left axis. The peak magnitudes of the strong El Niños (3-month Nino Region 3.4 average or ONI index) are indicated. The cgo-spo annual differences are plotted on a doubled right-hand scale. Dashed lines represent linear regressions through the annual average data.~~