Turnbull et al: Response to reviewers on revised manuscript

Formatted: Font:Bold

Comments from the editor and reviewers are in bold type, our responses are in plain type. The marked up manuscript and tables/figures are included in this document following the point-by-point response.

Editor: I have now considered the revised manuscript and all three reviewers. It is clear that the manuscript is much improved as a result of the revisions. However, the reviewers are still split in their opinions though noticeably less so than before. I would like you to address the points the reviewers raise, especially those by reviewers 2 and 3. If you wish to send a point-by-point response before revising the paper, that will be fine. If not, please submit your response with a revised manuscript.

To give you some guidance, I think some restructuring would make the paper much more readable. Section 3.5 could become a new section (4. Data validation), and 3.6 could form part of the new section 5.2 (Seasonal variability) which should lead to some reduction in overall length. I sympathise with reviewer 2's comment about the balance (and redundancy) between the main manuscript and the supplementary information. I would think that much of the curve fitting text would be better in the supplementary information as it underlies, rather than adds to, the main findings. It would make the paper more accessible to the non-specialist such as myself.

We changed Section 3.5 and it's subsections (3.5.1, 3.5.2 and 3.5.3) to Section 4 Data Validation with Subsections 4.1, 4.2 and 4.3 as suggested.

Moving section 3.6 Smooth Curve Fit to the results section was problematic, since we discuss the smooth curve in (renamed) Section 4 Data Validation. Therefore, we left this section in the methods (renumbered as section 3.5), but substantially shortened the main text and moved the rest to the supplementary material (supplementary section 5.4). We also made some minor changes in (renumbered) Section 5.2 Seasonal Variability in the Wellington Record to be consistent with the changes in the methods sections.

I also sympathise with reviewer 2's comment about Section 3.7 and the use of NAME as you do not refer to the analysis much in the text. I suggest zooming in on Figure 5. That could make it easier to assess the possibility of local contamination. If the suggested calculation can be done, that would be good. However I say that not knowing how hard it is to do.

We removed this section – we had considered doing so after the first round of reviews, and with these new reviews, we decided to indeed remove it. As for the suggested calculation, i is reasonably complex and time consuming, and given that the influence is quite small, we felt it was beyond the scope of this paper to add that analysis.

The other comments are more detailed and I look forward to your addressing them either in a revised manuscript or in your response.

Reviewer 1 (Samuel Hammer):

No changes suggested or made.

Reviewer 3 (John Miller):

a) Regarding gap-filling of time series for ccgcrv, there is a web site describing ccgcrv where one can also download all the code. www.esrl.noaa.gov/gmd/ccgg/mbl/crvfit/

Thank you, we added a link to the website in the text.

b) The hypothesis that a change in the land carbon sink could explain changes in the D14C latitude gradient is unrealistic considering that D14C is a quantity designed to factor out net fluxes from the land biosphere. line 647 of new pdf

We agree, and have added a statement to this sentence "although this latter is less likely since $\Delta^{14}CO_2$ is much less sensitive to biospheric fluxes than to either ocean or fossil fuel fluxes (e.g. Levin et al., 2010; Turnbull et al., 2009)."

c) WLG is still being used as a site code in Fig. 4.

Fixed. Thank you.

Reviewer 2:

We especially thank this reviewer for a very thorough review of the paper and some very insightful comments. We hope that the changes we have made have captured the reviewer's intent.

The authors have made some improvements and removed some unsupported statements in the revision. However, the abstract and conclusions are unchanged from the original version. These both include the misleading and unsupported statements about the Southern Ocean influence on the interhemispheric gradient noted in the first review, and they both need to be revised to remove these and to focus on the new results reported here.

We revised the abstract and conclusions to include the alternative (although less likely) explanations for the developing interhemispheric gradient.

The paper has a strange organization that mixes results (for comparison and validation of measurements) into the Methods section, and the Results and Discussion section is dominated by discussion (most of the data has been presented before and there is no new analysis of the causes of observed cycles, trend and gradient). The authors should revise the organization to have a Results section which is clearly showing new results from this analysis including the comparison and validation of measurements.

We reorganized the sections following the editor's recommendations – see comments above.

The Results section should clearly report revisions to the Wellington record that were made since Currie 2009 and new observations of the BHD D14C trend over 2005-14 (an interesting new result that is not currently included in the paper), in addition to the analysis of seasonal cycles and estimates of interhemispheric gradients with other stations already included.

We added a paragraph to section 5.1 discussing the recent BHD trend, and included a calculation showing that the Suess Effect is increasing and cannot explain the slowing in the BHD trend.

Then the Results and Discussion section should be renamed Discussion, and shortened.
We have followed the editor's suggestions for reorganization (see above).

Section 3.7: I still find this analysis and figure to not be very useful. Why don't the authors use the footprints to calculate fossil fuel CO2 from local combustion to make this point more clearly? In any case, please note here that the model simulations are only used to detect local influences of fossil fuel combustion and potential seasonality
We thought this through, and ultimately agreed that removing this information and figure was the best choice. Instead we have simply referenced Steinkamp et al., 2017, where the BHD footprints are discussed in some detail. We did consider using the footprints to calculate the fossil fuel CO_2 from local combustion, but this is a fairly major undertaking since it requires spatially and temporally explicit fossil fuel emissions for New Zealand. Since the influence is small, this appears to us excessive for this paper.

Section 4.2: This section is too long and it does not give a clear impression of the drivers of the seasonal cycle. Following the recommendation above for reorganization, the seasonal cycle results should be separated from the discussion, which can be shorted to a couple of paragraphs.
We extended this section in the first round of revisions, specifically to address questions that this reviewer raised. Shortening this section would necessitate removing this added discussion and we not done so.

Ocean exchange needs to be mentioned as an influence. The authors do not mention the analysis of Levin 2010 on seasonality at CGO, who found stratosphere-troposphere exchange and oceanic exchange were the main contributors.
Levin et al 2010 argued that seasonal variability in transport (cross-equator and strat-trop) were the main drivers and this is already stated in the text.

Also, Randerson 2002 Fig 6 for the 1980s.
We discuss the 1980s period and compare with the Randerson results in this section already. No changes.

The authors should be careful not to over-interpret the data in this section. Are the authors stating that the biosphere is the primary contributor to the seasonal cycle at BHD in the 1980s-90s when they say "This is consistent with a change in sign of the terrestrial biosphere contribution as the bomb 14C pulse began to return to the atmosphere from the biosphere (Randerson et al., 2002)"? Why would the biosphere contribution be strongest in winter?
We added some words to explain that the seasonality in atmospheric transport convolved with the flip in biosphere seasonality could drive the change.

L471 Also shown in Graven et al. 2012, La Jolla trend
Added this reference.

L578-580 It's strange to introduce Cape Grim here, when it has already been referred to several times and compared to BHD data in the previous section.
We added the Cape Grim info where Cape Grim is first discussed in the paper. We also left it here to help the reader follow where all the sites are.

This comparison with other sites should be introduced in the Methods rather than midway through the Results and Discussion. Did the authors consider laboratory offsets between the three laboratories?
We have opted to leave this section where it is, rather than adding another section in the methods. We did consider laboratory offsets and differences, and this is discussed briefly.

L593-601 This is not a new or robust estimate of the exchange time and needs to be deleted. Many studies have noted the delay of about a year between the peak D14C in the two hemispheres, e.g. Nydal 1966, with Lal and Rama 1966 providing a more rigorous estimate than offered here.
In preparing this paper, we were surprised how little published information there actually is on the use of the ^{14}C bomb spike to look at interhemispheric exchange time, since it is "common knowledge" that this is one of the applications of bomb ^{14}C measurements. Following this comment, we re-read those early papers, and they estimated exchange times of about 1.7 years. In neither case did they include data beyond 1964, which means that they did not capture the appearance of the bomb ^{14}C peak in the Southern Hemisphere.

L633 Levin et al. 2010 also attribute the change to the (Southern) Ocean
Agreed, and re-reading the text, we feel this is already stated sufficiently.

L636 Where is it shown that the interhemispheric gradient develops rapidly? From Levin 2010, Fig 6i there appears to be a clear trend over 1985-2005, albeit with interannual variation. No such figure is included here.
We added a reference to figure 5, where this is shown. Note also that the development and continuation of this gradient is clearer with the longer records shown here than in either the Graven or Levin papers.

L638-651 should be deleted and replaced with one sentence such as "The interhemispheric gradient may also be influenced by the apparent reorganization of Southern Ocean carbon exchange in the early 2000s (Landschützer et al., 2015), which is postulated to be associated with changes in upwelling of deep water (DeVries et al., 2017) to which atmospheric Δ14CO2 is highly sensitive (Rodgers et al., 2011; Graven et al., 2012b)."
We disagree on this point. The most likely explanation for the interhemispheric ^{14}C gradient is Southern Ocean upwelling, and there is wide interest in the possibilities of

using ^{14}C to understand Southern Ocean carbon cycling. Thus we think it deserves more than one sentence to explain the rationale.

For "aging", it may be better to say "dating"
We think the reviewer is referring to "aging" of recent materials. We have retained the word "aging" since "dating" is not normally used for post-bomb radiocarbon age determination.

The authors argue for retaining the data with large scatter over 1995-2005, but these data flagged "T" and the smoothed record over this period need to at least include larger uncertainties. The two smoothed records provided "BHD" and "BHDCGO" differ by up to 10 per mil, but the reported uncertainties are about 2 per mil. Shouldn't the RMSE of 5 per mil over this period be incorporated in the uncertainty of the smoothed record of BHDCGO?
There are many possible strategies for evaluating uncertainties. Our view is that for this work it is most appropriate to report uncertainties based on the dataset itself, rather than attempting to incorporate biases diagnosed from other datasets. The atmospheric radiocarbon community is in the process of strengthening intercomparisons between laboratories and developing harmonized global and regional datasets. Uncertainty analysis from comparison of multiple records such as that suggested here will be incorporated into that future work.

Can more information be provided on the "additional error term, determined from the long-term repeatability of secondary standard materials"? The Turnbull 2016 reference appears to report only recent secondary standard measurements since 2011.
This additional error term has been added to measurements for all Rafter AMS measurements, but the methodology had never been published prior to the Turnbull et al 2015 paper. A longer discussion is included in the supplementary material and we added a reference to that in the main text.

The excel spreadsheet refers to WLG in the header text on several sheets.
Fixed. Thank you.

It seems unnecessary to repeat so much text from the paper in the supplement.
The supplement includes more detail of the dataset than is in the paper, and we feel that it becomes difficult to follow if users need to jump between the main paper and the supplement to get all the information, therefore we prefer to include all the detail in the supplement.

Supplement does not define "trend" and "fit" reported in the spreadsheet. "Trend" suggests a rate of change rather than deseasonalized data.
We moved the bulk of the information on the smooth curve fits from the main paper into the supplementary material and added into comments there to indicate how "trend" and "fit" are defined in the datasets.

Sixty years of radiocarbon dioxide measurements at Wellington, New Zealand 1954 – 2014

Jocelyn C. Turnbull[1,2,*] Sara E. Mikaloff Fletcher[3], India Ansell[1], Gordon Brailsford[3], Rowena Moss[3], Margaret Norris[1], Kay Steinkamp[3]

[1]GNS Science, Rafter Radiocarbon Laboratory, Lower Hutt, New Zealand
[2]CIRES, University of Colorado at Boulder, Boulder, Colorado, USA
[3]NIWA, Wellington, New Zealand
* contact author: j.turnbull@gns.cri.nz

1. Abstract

We present 60 years of $\Delta^{14}CO_2$ measurements from Wellington, New Zealand (41°S, 175°E). The record has been extended and fully revised. New measurements have been used to evaluate the existing record and to replace original measurements where warranted. This is the earliest direct atmospheric $\Delta^{14}CO_2$ record and records the rise of the ^{14}C "bomb spike", the subsequent decline in $\Delta^{14}CO_2$ as bomb ^{14}C moved throughout the carbon cycle and increasing fossil fuel CO_2 emissions further decreased atmospheric $\Delta^{14}CO_2$. The initially large seasonal cycle in the 1960s reduces in amplitude and eventually reverses in phase, resulting in a small seasonal cycle of about 2 ‰ in the 2000s. The seasonal cycle at Wellington is dominated by the seasonality of cross-tropopause transport, and differs slightly from that at Cape Grim, Australia, which is influenced by anthropogenic sources in winter. $\Delta^{14}CO_2$ at Cape Grim and Wellington show very similar trends, with significant differences only during periods of known measurement uncertainty. In contrast, similar clean air sites in Northern Hemisphere show a higher and earlier bomb ^{14}C peak, consistent with a 1.4-year interhemispheric exchange time. From the 1970s until the early 2000s, the Northern and Southern Hemisphere $\Delta^{14}CO_2$ were quite similar, apparently due to the balance of ^{14}C-free fossil fuel CO_2 emissions in the north and ^{14}C-depleted ocean upwelling in the south. The Southern Hemisphere sites show a consistent and marked elevation above the Northern Hemisphere sites since the early 2000s, which is most likely due to reduced upwelling of ^{14}C-depleted and carbon-rich deep waters in the Southern Ocean, although an underestimate of fossil fuel CO_2 emissions or changes in biospheric exchange are also possible explanations. This developing $\Delta^{14}CO_2$ interhemispheric gradient is consistent with recent studies that indicate a reinvigorated Southern Ocean carbon sink since the mid-2000s, and suggests that upwelling of deep waters plays an important role in this change.

2. Introduction

Measurements of radiocarbon in atmospheric carbon dioxide ($\Delta^{14}CO_2$) have long been used as a key to understanding the global carbon cycle. The first atmospheric $\Delta^{14}CO_2$ measurements were begun at Wellington, New Zealand in 1954 (Rafter, 1955; Rafter et al., 1959), aiming to better understand carbon exchange processes (Otago Daily Times, 1957). Northern Hemisphere $\Delta^{14}CO_2$ measurements began a few years later in 1962, in Norway (Nydal and Løvseth, 1983) and 1959 in Austria (Levin et al., 1985).

^{14}C is a cosmogenic nuclide produced naturally in the upper atmosphere through neutron spallation, reacts rapidly to form ^{14}CO and then oxidizes to $^{14}CO_2$ over a period of 1- 2 months, after which it moves throughout the global carbon cycle. Natural ^{14}C production is roughly balanced by radioactive decay, which mostly occurs in the carbon-rich and slowly overturning ocean carbon reservoir and to a lesser extent in the faster cycling terrestrial carbon reservoir. The perturbations to $\Delta^{14}CO_2$ from atmospheric nuclear weapons testing in the mid-20th century and additions of ^{14}C-free CO_2 from fossil fuel burning have both provided tools to investigate CO_2 sources and sinks.

Penetration of bomb-^{14}C into the oceans has been used to understand ocean carbon uptake processes (Oeschger et al., 1975; Broecker et al., 1985; Key et al., 2004; Naegler et al., 2006; Sweeney et al., 2007). Terrestrial biosphere carbon residence times and exchange processes have also been widely investigated using bomb-^{14}C (e.g. Trumbore et al., 2000; Naegler et al., 2009). Stratospheric residence times, cross-tropopause transport and interhemispheric exchange can also be examined with atmospheric $\Delta^{14}CO_2$ observations (Kjellström et al., 2000; Kanu et al., 2015).

The Suess Effect, the decrease in atmospheric $\Delta^{14}CO_2$ due to the addition of ^{14}C-free fossil fuel CO_2, was first identified in 1955 (Suess, 1955). It has subsequently been refined (Meijer et al., 1996; Levin et al., 2003; Turnbull et al., 2006) and used to investigate fossil fuel CO_2 additions at various scales (e.g. Turnbull et al., 2009a; Djuricin et al., 2010; Miller et al., 2012; Lopez et al., 2013; Turnbull et al., 2015).

The full atmospheric ^{14}C budget has been investigated using long term $\Delta^{14}CO_2$ records in conjunction with atmospheric transport models (Caldiera et al., 1998; Randerson et al., 2002; Naegler et al., 2006; Turnbull et al., 2009b; Levin et al., 2010). These have shown changing controls on $\Delta^{14}CO_2$ through time. Prior to nuclear weapons testing, natural cosmogenic production added ^{14}C in the upper atmosphere, which reacted to CO_2 and moved throughout the atmosphere and the carbon cycle. The short carbon residence time in the biosphere meant that biospheric exchange processes had only a small influence on $\Delta^{14}CO_2$, whereas the ocean exerted a stronger influence due to radioactive decay during its much longer (and temporally varying) turnover time. The addition of bomb ^{14}C in the 1950s and 1960s almost doubled the atmospheric ^{14}C content. This meant that both the ocean and biosphere were very ^{14}C-poor relative to the atmosphere in the two decades following the atmospheric test ban treaty. As the bomb-^{14}C was distributed throughout the carbon cycle, this impact weakened, and by the 1990s, the additions of fossil fuel CO_2 became the largest contributor to the $\Delta^{14}CO_2$ trend (Randerson et al., 2002; Turnbull et al., 2007; Levin et al., 2010; Graven et al., 2012).

The long-term $\Delta^{14}CO_2$ records have been crucial in all of these findings, and the Wellington $\Delta^{14}CO_2$ record is of special importance, being the oldest direct atmospheric trace gas record, even predating the CO_2 mole fraction record started at Mauna Loa in 1958 (Keeling, 1961; Keeling and Whorf, 2005). It is the only Southern Hemisphere record recording the bomb spike. Several short Southern Hemisphere records do exist (Manning et al., 1990; Meijer et al., 2006; Graven et al., 2012b; Hua and Barbetti, 2013), and some longer records began in the 1980s (Levin et al., 2010). Over the more than 60 years of measurement, there have necessarily been changes in how the Wellington samples are collected and measured. There are no comparable records during the first 30 years of measurement, so that the data quality has not been independently evaluated. Comparison with other records since the mid-1980s has suggested that there may be biases in some parts of the Wellington record (Currie et al., 2011).

Here we present a revised and extended Wellington atmospheric $^{14}CO_2$ record, spanning 60 years from December 1954 to December 2014. We detail the different sampling, preparation and measurement techniques used through the record, compare with new tree ring measurements, discuss revisions to the previously published data and provide a final dataset with accompanying smooth curve fit.

In the results and discussion, we revisit the key findings that the Wellington $^{14}CO_2$ record has provided over the years and expand with new findings based on the most recent part of the record. The most recent publication of this dataset included data to 2005 (Currie et al., 2011) and showed periods of variability and a seasonal cycle at Wellington that differ markedly from the independent Cape Grim, Tasmania $^{14}CO_2$ record at a similar southern latitude (Levin et al., 2010). Here we add complementary new data to investigate these differences, fill gaps and extend the record to near-present.

3. Methods

Over 60 years of measurement, a number of different sample collection, preparation, measurement and reporting methods have been used. In this section, we give an overview of the various methods and changes through time, and they are summarized in table 1. Full details of the sampling methods used through time are provided in the supplementary material, compiling methodological information documented in previous reports on the Wellington record (Rafter and Fergusson, 1959; Manning et al., 1990; Currie et al., 2011) along with methods newly applied in this new extension and refinement of the dataset.

3.1. Sampling sites

Samples from 15 December 1954 – 5 June 1987 were collected at Makara (Lowe, 1974), on the south-west coast of the North Island of New Zealand (MAK, 41.25°S, 174.69°E, 300 m asl). Samples since 8 July 1988 have been collected at Baring Head (Brailsford et al., 2012) on the South Coast of the lower North Island and 23 km southeast of Makara (BHD, 41.41°S, 174.87°E, 80 m asl) (figure 1). We also discuss tree ring samples collected from Eastbourne, 12 km north of Baring Head on Wellington Harbour.

3.2. Collection methods

3.2.1. NaOH absorption

The primary collection method is static absorption of CO_2 into nominally CO_2-free 0.5 or 1 M sodium hydroxide (NaOH) solution, which is left exposed to air at the sampling site providing an integrated sample over a period of ~2 weeks (section S3.1; Rafter, 1955). From 1954-1995, ~ 2 L NaOH solution was exposed to air in a large (~450 cm^2 surface area) Pyrex® tray. Since 1995, wide-mouth high-density polyethylene (HDPE) bottles containing ~200 mL NaOH solution were left open inside a Stevenson meteorological screen; the depth of the solution in the bottles remained the same as that in the previously used trays. No significant difference has been observed between the two methods (Currie et al., 2011). A few early (1954-1970) samples were collected using different vessels, air pumped through the NaOH (vs. passive absorption), or NaOH was replaced with barium hydroxide (Rafter, 1955; Manning et al., 1990). CO_2 is extracted from the NaOH solution by acidification followed by cryogenic distillation (Rafter and Fergusson, 1959; Currie et al., 2011). Static NaOH absorption necessarily fractionates relative to CO_2 in the atmosphere. Typical $\delta^{13}C$ values are -15 to -25 ‰ for these samples, and this is corrected for in the data analysis.

3.2.2. Whole air flasks

In this study, we use whole air flask samples collected at Baring Head to supplement and/or replace NaOH samples. Flasks of whole air are collected by flushing ambient air through the flask for several minutes then filled to slightly over ambient pressure. Most flasks were collected during southerly, clean air conditions (Stephens et al., 2013). CO_2 is extracted cryogenically (Turnbull et al., 2015). For whole air samples collected from 1984-1993, the extracted CO_2 was archived until 2012. We evaluated the quality of this archived CO_2 using two methods. Tubes with major leakage were readily detected by air present in the tube and were discarded. $\delta^{13}C$ from all the remaining samples was in agreement with $\delta^{13}C$ measured from separate flasks collected at Baring Head and measured for $\delta^{13}C$ by Scripps Institution of Oceanography at close to the time of collection (http://scrippsco2.ucsd.edu/data/nzd). Whole air samples collected since 2013 are analyzed for $\delta^{13}C$ and other trace gases and isotopes at NIWA (Ferretti et al., 2000) and for the $^{14}CO_2$ measurement, CO_2 is extracted from whole air at Rafter Radiocarbon Laboratory (Turnbull et al., 2015).

3.2.3. Tree rings

When trees photosynthesize, they faithfully record the $\Delta^{14}C$ of ambient CO_2 in their cellulose, the structural component of wood. Annual tree rings therefore provide a summertime (approximately September – April in the Southern Hemisphere) daytime average $\Delta^{14}CO_2$. Photosynthetic uptake varies during the daylight hours depending on factors including growth period, sunlight, and temperature (Bozhinova et al., 2013), resulting in a somewhat different effective sampling pattern than the 1-2 week NaOH solution collections. We show in section 3.5.1. that at the Wellington location this difference is negligible. Note that we assign the mean age of each ring as January 1 of the year in which growth finished (i.e. the mean age of a ring growing from September – April), whereas dendrochronologists assign the "ring year" as the year in which ring growth started (i.e. the previous year).

We collected cores from three trees close to the Baring Head site. A pine (*Pinus radiata*) located 10 m from the Baring Head sampling station (figure 1) yielded rings back to 1986 (Norris, 2015). A longer record was obtained from two New Zealand kauri (*Agathis australis*) specimens planted in 1919 and 1920, located 20 m from one another in Eastbourne, 12 km from Baring Head (figure 1). Kauri is a long-lived high-density softwood species that has been widely used in dendrochronology and radiocarbon calibration studies (e.g. Hogg et al., 2013).

Annual rings were counted from each core. Shifting the Eastbourne record by one year in either direction moves the ^{14}C bomb spike maximum out of phase with the NaOH-based Wellington $\Delta^{14}CO_2$ record (supplementary figure S1), confirming that the ring counts are correct. For the Baring Head pine, rings go back to only 1986, and we verify them by comparing with the Eastbourne record. They show an insignificant mean difference of -0.4 ± 0.8 ‰ (supplementary figure S1).

In practice, it is difficult to ensure that one annual ring is sampled without losing any material from that ring, and no wood from surrounding rings is included. To evaluate the potential bias from this source, we measured replicate samples from different cores from the same tree (Baring Head) or two different trees (Eastbourne). For samples collected since 1985, all these replicates are consistent within their assigned uncertainties (supplementary figure S2). However, for three replicates from Eastbourne in 1963, 1965 and 1971, we see large differences of 9.2, 44.5 and 4.9 ‰, which we attribute to small differences in sampling of the rings that were magnified by the rapid change in Δ^{14}C of up to 200 ‰ yr^{-1} during this period. Thus, the tree ring Δ^{14}C values during this period should be treated with caution.

Cellulose was isolated from whole tree rings by first removing labile organics with solvent washes, then oxidation to isolate the cellulose from other materials (Norris, 2015; Hua et al., 2000). The cellulose was combusted and the CO_2 purified following standard methods in the Rafter Radiocarbon Laboratory (Baisden et al., 2013).

3.3. ^{14}C measurement

Static NaOH samples were measured by conventional decay counting on the CO_2 gas from 1954 – 1995 (Manning et al., 1990; Currie et al., 2011) and these are identified by their unique "NZ" numbers. All measurements made since 1995, including recent measurements of flask samples collected in the 1980s and 1990s, were reduced to graphite, measured by accelerator mass spectrometry (AMS), and are identified by their unique "NZA" numbers. The LG1 graphitization system was used from 1995 to 2011 (NZA < 50,000) (Lowe et al., 1987), and replaced with the RG20 graphite system in 2011 (NZA > 50,000) (Turnbull et al., 2015). Samples measured by AMS were stored for up to three years between sample collection and extraction/graphitization/measurement.

For samples collected from 1995 to 2010, an EN Tandem AMS was used for measurement (NZA < 35,000, Zondervan and Sparks, 1996). Until 2005 (NZA <30,000, including all previously reported Wellington $^{14}CO_2$ data), only ^{13}C and ^{14}C were measured on the EN Tandem system, so the normalization correction for isotopic fractionation (Stuiver and Polach, 1977) was performed using an offline isotope ratio mass spectrometer $\delta^{13}C$ value. The data reported from 2005 onwards (NZA > 30,000) show a reduction in scatter reflecting the addition of online ^{12}C measurement in the EN Tandem system in 2005. This allows direct online correction for isotopic fractionation that may occur during sample preparation and in the AMS system (Zondervan et al., 2015), and results in improved long-term repeatability. Fractionation in the AMS system may vary in sign depending on the particular conditions, but incomplete graphitization biases the graphite towards lighter isotopes, which, if undiagnosed, will bias $\Delta^{14}C$ high. The LG1 graphitisation system used during this period did not directly evaluate whether graphitization was complete, so it is possible or even likely that there was a high bias in the 1995 – 2005 measurements. This is further discussed in section 3.5.3.

For all EN Tandem samples, a single large aliquot of extracted CO_2 was split into four separately graphitized and measured targets and the results of all four were averaged. We have revisited the multi-target averaging, applying a consistent criterion to exclude outliers and using a weighted mean of the retained measurements (supplementary material). This results in differences of up to 5 ‰ relative to the values reported by Currie et al. (2011) and is discussed in more detail in the supplementary material.

In 2010, the EN Tandem was replaced with a National Electrostatics Corporation AMS, dubbed XCAMS (NZA > 34,000). XCAMS measures all three carbon isotopes, such that the normalization correction is performed using the AMS measured ^{13}C values (Zondervan et al., 2015). XCAMS measurements are made on single graphite targets measured to high precision of typically 1.8 ‰ (Turnbull et al., 2015).

3.4. Results format

NaOH samples are collected over a period of typically two weeks, and sometimes much longer. We report the date of collection as the average of the start and end dates. In cases where the end date was not recorded, we use the start date. For a few samples, the sampling dates were not recorded or are ambiguous, and those results have been excluded from the reported dataset.

Results are reported here as $F^{14}C$ (Reimer et al., 2004) and $\Delta^{14}C$ (Turnbull et al., 2007). $F^{14}C$ is corrected for isotopic fractionation and blank corrected. We calculated $F^{14}C$ from the original measurement data recorded in our databases, and updated a handful of records where transcription errors were found. $\Delta^{14}C$ is derived from $F^{14}C$, and corrected for radioactive decay since the time of collection; this is slightly different from $\Delta^{14}C$ as defined by Stuiver and Polach (1977) that is corrected to the date of measurement. $\Delta^{14}C$ has been recalculated using the date of collection for all results, resulting in changes of a few tenths of permil in most $\Delta^{14}C$ values relative to those reported by Currie et al. (2011) and Manning et al. (1990). Uncertainties are reported based on the counting statistical uncertainty and for AMS measurements we add an additional error term, determined from the long-term repeatability of secondary standard materials (Turnbull et al., 2015; supplementary section 5.3). Samples for which changes have been made relative to the previously published results are indicated by the quality flag provided in the supplementary dataset. Where more than one measurement was made for a given date, we report the weighted mean (Bevington and Robinson, 2003) of all measurements.

3.5. Smooth curve fits

In addition to the raw measured $\Delta^{14}CO_2$ values, we calculate a smooth curve fit and deseasonalized trend from the Wellington $\Delta^{14}C$ and $F^{14}C$ datasets. The deseasonalized trend may be more useful than the raw data for aging of recent materials (e.g. Reimer et al., 2004; Hua et al., 2013). Acknowledging that the 1995-2005 period is variable and possibly biased in the Wellington record (section 4.3), we also provide in the supplementary material an alternative mid-latitude Southern Hemisphere smooth curve fit and deseasonalized trend in which the Wellington data for 1995-2005 has been removed and replaced with the Cape Grim, Tasmania data for that period (Levin et al., 2010).

Curvefitting is particularly challenging for the $\Delta^{14}CO_2$ record, since (a) there are data gaps and inconsistent sampling frequency, (b) the growth rate and trend vary dramatically and (c) the seasonal cycle changes both in magnitude and phase (Section 5.2). We chose to use the CCGCRV fitting procedure (Thoning et al., 1989; www.esrl.noaa.gov/gmd/ccgg/mbl/crvfit/), which can readily handle the data gaps, inconsistent sampling frequency, and rapid changes in the trend. To address the changing seasonal cycle, we make separate fits to the record for five time periods: 1954-1965, 1966-1979, 1980-1989, 1990-2004, 2005-2014. These divisions were chosen based on major changes in the raw observational growth rate, seasonal cycle and data quality (supplementary section 6). For each time period, we use CCGCRV with one linear and two harmonic terms and fit residuals are added back using a low-pass filter with an 80-day cutoff in the frequency domain. At each transition, we overlapped a two-year period and linearly interpolated the two fits across that two-year period to smooth the transitions caused by end effects. The one-sigma uncertainty on the smoothed curve and deseasonalized trend were determined using a Monte Carlo technique. Further details of the fitting procedure and choice of time period cutoffs are provided in the supplementary material.

The mean difference between the fitted curve and the measured $\Delta^{14}CO_2$ values is 3.8 ‰, consistent with the typical measurement uncertainty for the full dataset. Further, the residuals are highest for the early period (1954-1970) at 6 ‰, consistent with the larger measurement errors at that time of ~6 ‰. The residuals improve as the measurement errors reduce, such that since 2005, the mean residual is 2 ‰, consistent with the reported 2 ‰ uncertainties. The exception is the 1995- 2005 period where the mean residual difference of 5 ‰ is substantially higher than the mean reported uncertainty of 2.5 ‰, reflecting the apparent larger scatter during this period (Section 4.3).

4. Data validation

Deleted: 3.5

Formatted: Heading 1

4.1. Tree ring comparison

Deleted: 3.5

Over the more than 60 years of the Wellington $\Delta^{14}CO_2$ record, there have necessarily been many changes in methodology, and the tree rings provide a way to validate the full record, albeit with lower resolution. Due to the possible sampling biases in the tree rings (section 3.2.3.), we do not include them in the final updated record, but use them to validate the existing measurements.

During the rapid $\Delta^{14}CO_2$ change in the early 1960s, there are some differences between the kauri tree ring and Wellington $\Delta^{14}CO_2$ records (Figure 2). The 1963 and 1964 tree ring samples are slightly lower than the concurrent $\Delta^{14}CO_2$ samples. The peak $\Delta^{14}CO_2$ measurement in the tree rings is 30 ‰ lower than the smoothed $\Delta^{14}CO_2$ record, and 100‰ lower than the two highest $\Delta^{14}CO_2$ measurements in 1965. These differences are likely due to small errors in sampling of the rings, which will be most apparent during periods of rapid change.

Prior to 1960 and from the peak of the bomb spike in 1965 until 1990, there is remarkable agreement between the tree rings and Wellington $\Delta^{14}CO_2$ record, with the variability replicated in both records. And since 2005, there is excellent agreement across all the different records. Some differences are observed in 1990-1993 and 1995-2005, which we discuss in the following sections.

4.2. 1990-1993 anomaly

An anomaly in the gas counting measurements between 1990 and 1993 has previously been noted (figures 2, 3) as a deviation from the Cape Grim $\Delta^{14}CO_2$ record (CGO, 40.68°S, 144.68°E, 94 m asl, Levin et al., 2010) during the same period. Cape Grim is at similar latitude, and observes a mixture of air from the mid-latitude Southern Ocean sector and mainland Australia (Ziehn et al., 2014; Law et al., 2010). The Wellington and Cape Grim records overlap during almost all other periods (figure 3).

We use archived CO_2 from flask samples to evaluate this period of deviation. First, the recent flask samples collected since 2013 (n=12) agree very well with the NaOH static samples from the same period (figure 2), indicating that despite the difference in sampling period for the two methods, flask samples reflect the $\Delta^{14}CO_2$ observed in the longer-term NaOH static samples. We then selected a subset of archived 1984 - 1992 extracted CO_2 samples for measurement, mostly from Southerly wind conditions, but including a few from other wind conditions. These flask $\Delta^{14}CO_2$ measurements do not exhibit the anomaly seen in the NaOH static samples (figure 2), implying that the deviation observed in the original NaOH static samples may be a consequence of sampling, storage or measurement errors. Annual tree rings from both the kauri and pine follow the flask measurements for this period (figure 2), confirming that the NaOH static samples are anomalous.

The 1990-1993 period was characterized by major changes in New Zealand science, both in the organizational structure and personnel. Although we are unable to exactly reconstruct events at that time, we hypothesize that the NaOH solution was prepared slightly differently, perhaps omitting the barium chloride precipitation step for these samples. This would result in contaminating CO_2 absorbed on the NaOH before the solution was prepared. Since atmospheric $\Delta^{14}CO_2$ is declining, this would result in higher $\Delta^{14}CO_2$ observed in these samples than in the ambient air. Another possibility is that there were known issues with the background contamination in the proportional

Deleted: variability replicated

Deleted: 3.5

Formatted: Heading 2

counters during this period that could result in a high bias $\Delta^{14}CO_2$. In any case, these values are anomalous and we remove the original NaOH static sample measurements between 1990 and 1993 and replace them with the new flask measurements for the same period.

4.3. 1995-2005 variability

As already discussed in section 3.3, the measurement method was changed from gas counting to AMS for samples collected in 1995 and thereafter. During the first ten years of AMS measurements, the record is much noisier than during any other period (figure 2). Until 2005, offline $\delta^{13}C$ measurements on the evolved CO_2 were used in the normalization correction. In 2005, online ^{12}C measurement was added to the AMS system, allowing online AMS measurement of the $\delta^{13}C$ value and accounting for any fractionation during sample preparation and AMS measurement (Zondervan et al., 2015; see also section 3.3). This substantially improved the measurement accuracy and the noise in the $\Delta^{14}CO_2$ record immediately reduced as can be seen in the lower panel of figure 2. Therefore, we suspect that the variability and apparent high bias in the 1995-2005 period of the $\Delta^{14}CO_2$ record is due to measurement uncertainty and bias rather than atmospheric variability.

The remaining NaOH solution for all samples collected since 1995 has been archived, and typically only every second sample collected was measured, with the remainder archived without extraction. In 2011-2016, we revisited the 1995-2005 period, remeasuring some samples that had previously been measured and some that had never been measured for a total of 52 new analyses.

The new measurements for this time period do show reduced scatter over the original analyses, particularly for the period from 1998-2001 where the original analyses appear anomalously low and in 2002-2003 when the original analyses appear anomalously high. Yet there remain a number of both low and high outliers in the new measurements. These are present in both the samples that were remeasured and in those for which this was the first extraction of the sample. This suggests that a subset of the archived sample bottles were either contaminated at the time of collection, or that some bottles were insufficiently sealed, causing contamination with more recent CO_2 during storage. Comparison with the tree ring measurements and with the Cape Grim record (Levin et al., 2010) suggest that the measurements during this period may, on average, be biased high as well as having additional scatter (figure 3). Nonetheless, in the absence of better data, we retain both the original and remeasured NaOH sample results in the full Wellington record, with a special flag to allow users to easily remove the questionable results if they prefer. We also provide a smoothed fit that excludes these data (section 3.6).

5. Results and Discussion

5.1. Variability in the Wellington record through time

The Wellington $\Delta^{14}CO_2$ record begins in December 1954, at a roughly pre-industrial $\Delta^{14}CO_2$ level of -20 ‰ (figure 2). From 1955, $\Delta^{14}CO_2$ increased rapidly, near doubling

Deleted: 3.5

Formatted: Heading 2

Deleted: Therefore

Deleted:
3.7 Atmospheric Model Simulations
4

Deleted: 3.7 Atmospheric Model Simulations
4

Formatted: Heading 1

Deleted: 4

to 700 ‰ in 1965 at Wellington, due to the production of ^{14}C during atmospheric nuclear weapons tests. Nuclear tests in the early 1950s contributed to the rise, then a hiatus in testing in the late 1950s led to a plateau in Wellington $\Delta^{14}CO_2$ before a series of very large atmospheric tests in the early 1960s led to further increases (Rafter and Ferguson, 1959; Manning et al., 1990).

Most atmospheric nuclear weapons testing ceased in 1963, and the Wellington $\Delta^{14}CO_2$ record peaks in 1965 then begins to decline, at first rapidly at -30 ‰ yr^{-1} in the 1970s and gradually slowing to -5 ‰ yr^{-1} after 2005. The initial rapid decline has been attributed primarily to the uptake of the excess radiocarbon into the oceans, and to a lesser extent, uptake into the terrestrial biosphere (Naegler et al 2006; Randerson et al., 2002; Manning et al., 1990; Stuiver and Quay 1981). The short residence time of carbon in the biosphere means that from the 1980s, the terrestrial biosphere changed from a ^{14}C sink to a ^{14}C source as the bomb pulse was re-released (Randerson et al., 2002; Levin et al., 2010).

Natural cosmogenic production of ^{14}C damps the rate of decline since the bomb peak by ~5 ‰ yr^{-1} in $\Delta^{14}CO_2$; this may vary with the solar cycle, but there is no known long-term trend in this component of the signal (Turnbull et al., 2009; Naegler et al., 2006). There is also a small positive contribution from the nuclear industry which emits ^{14}C to the atmosphere, and this has increased from zero in the 1950s to 0.5 – 1 ‰ yr^{-1} in the last decade (Turnbull et al., 2009b; Levin et al., 2010; Graven and Gruber, 2011).

The Suess Effect, the decrease in atmospheric $\Delta^{14}CO_2$ due to the addition of ^{14}C-free fossil fuel CO_2 to the atmosphere (Suess, 1955; Tans, 1979; Levin et al., 2003), was first recognized in 1955 and has played a role throughout the record. Although the magnitude of fossil fuel CO_2 emissions has grown through time, when convolved with the declining atmospheric $\Delta^{14}CO_2$ history, the impact on $\Delta^{14}CO_2$ stayed roughly constant at -10 ‰ yr^{-1} from the 1970s to the mid-2000s (Randerson et al., 2002; Levin et al., 2010; Graven et al., 2012). Yet the continued increase in fossil fuel CO_2 emissions has slightly increased the impact of fossil fuel CO_2 in the last few years, to about -12 ‰ yr^{-1} in 2014 (using annual global fossil fuel CO_2 estimates from CDIAC (Boden et al., 2017)). Since the 1990s, the Suess Effect has been the largest driver of the ongoing negative growth rate (Turnbull et al., 2009b; Levin et al., 2010).

The most recent part of the Wellington $\Delta^{14}CO_2$ record from 2005 – 2014 is reported here for the first time. It shows a continuing downward trend in $\Delta^{14}CO_2$ of -5 ‰ yr^{-1}, a slight slowing in the negative trend relative to the 1990 – 2004 period which had a trend of 5.8 ‰ yr^{-1}. This slight slowing in the downward $\Delta^{14}CO_2$ trend is opposite of what might be expected due to the Suess Effect alone. Possible explanations are a slowing of the rate of uptake of ^{14}C into the oceans, an increase in the return rate of bomb ^{14}C to the atmosphere from the biosphere, and a secular increase in ^{14}C production.

5.2 Seasonal variability in the Wellington record

We determine the changing seasonal cycle from smooth curve fits to five separate periods of the record (1954-1965, 1966-1979, 1980-1989, 1990-2004, 2005-2014, figure 4 top panel). This subdivision is necessary to allow the seasonal cycle to vary through time

Deleted: 2003)

Deleted: has

Deleted: yr^{-1} since

Deleted:).

Formatted: Superscript

Deleted: 4

Deleted: .

since the CCGCRV curve fitting routine assigns a single set of harmonics to the time period fitted (see section 3.5). The choice of time periods is discussed in supplementary section 6. We also created detrended, fitted $\Delta^{14}CO_2$ seasonal cycles by subtracting the deseasonalised trend from the observations. Comparison with the detrended, fitted seasonal cycle determined from the smooth curve fits (figure 4 bottom panel) shows that the smooth curve fit, as might be expected, does not capture the largest deviations from the trend seen in the observations, but represents the changing seasonal cycle quite well.

Deleted: 6

Deleted: values

The 1966-1979 period shows a strong seasonal cycle (figure 4) with a consistent phase and an amplitude that varies from a maximum in 1966 of 30‰ gradually declining to 3 ‰ in 1979, with a mean amplitude of about 6 ‰. This is primarily attributed to seasonally varying stratosphere – troposphere exchange bringing bomb ^{14}C into the troposphere (Manning et al., 1990; Randerson et al., 2002). Manning et al. (1990) were unable to simulate the correct phasing of the seasonal cycle, apparently because their model distributed bomb ^{14}C production throughout both Northern and Southern stratosphere. In fact, the majority of the bomb ^{14}C was produced in the Northern Hemisphere stratosphere (Enting et al., 1982). Randerson et al (2002) were able to match the amplitude of the Wellington seasonal cycle during this time period, although their model was out of phase with the observations by about 1.5 months. They attribute the seasonal cycle during this period mostly to the seasonality in Northern Hemisphere stratosphere – troposphere exchange with a phase lag caused by cross-equator exchange into the Southern Hemisphere. The seasonal cycle kept the same phase but gradually decreased in amplitude until the late 1970s, attributed to the declining disequilibrium between the stratosphere and troposphere as the bomb ^{14}C moved throughout the carbon reservoirs.

Between 1978 and 1980 the seasonal cycle weakened, and then reversed during the 1980s, with a maximum in winter (June – August) and amplitude of about 2 ‰. The detrended observations show that this change in phase is not an artifact of the fitting method (bottom panel of figure 4). This result is comparable to that obtained by Manning et al. (1990) and Currie et al. (2011), who both used a seasonal trend loess (STL) procedure to determine the seasonal cycle from the same data. This is consistent with the seasonality in atmospheric transport convolving with a change in sign of the terrestrial biosphere contribution as the bomb ^{14}C pulse began to return to the atmosphere from the biosphere (Randerson et al., 2002).

Deleted: .

The Wellington $\Delta^{14}CO_2$ seasonal cycle declined in the 1990s, and the larger variability in the observations between 1995 and 2005 makes it difficult to discern a seasonal cycle during that period. Since 2005, the more precise measurements allow us to detect a small seasonal cycle with amplitude of about 2 ‰ (figure 4). We compare the seasonal cycle at Wellington from 2005 – 2015 with the seasonal cycle at Cape Grim, Australia from 1995-2010. There is no significant difference in the seasonal cycle at either site if we select only the overlapping time period of 2005-2010. Both sites show a similar magnitude seasonal cycle during this period, and Cape Grim shows a maximum in March – April that has been attributed primarily to the seasonality of atmospheric transport of Northern Hemisphere fossil fuel emissions to the Southern troposphere (Levin et al., 2010). This maxima at Cape Grim coincides with a seasonal maximum in the Wellington record. However, Wellington $\Delta^{14}CO_2$ exhibits a second maximum in the austral spring (October) that is not apparent at Cape Grim.

Recent work has shown that during the winter, the Cape Grim station is influenced by air coming off the Australian mainland including the city of Melbourne (Ziehn et al., 2014), which would act to reduce $\Delta^{14}CO_2$ at Cape Grim relative to Southern Ocean clean air. This shift is shown to be the result of seasonal variations in atmospheric transport. The two-week integrated sampling used for $\Delta^{14}CO_2$ at both Cape Grim and Baring Head means that in contrast to other species, $\Delta^{14}CO_2$ measurements cannot be screened to remove these pollution events.

In contrast, the Baring Head location near Wellington does not show significant seasonal variation in atmospheric transport (Steinkamp et al., 2017) and Baring Head is less likely than Cape Grim to be influenced by anthropogenic emissions in any season. Air is typically from the ocean, and the local geography means that the urban emission plume from Wellington and its northern suburbs of Lower Hutt very rarely passes over Baring Head (figure 1) and the typically high wind speeds further reduce the influence of the local urban area (Stephens et al., 2013). During the austral autumn, there is some land influence from the Christchurch region in the South Island, but emissions from Christchurch are much smaller than the Melbourne emissions influencing Cape Grim: State of Victoria fossil fuel CO_2 emissions for 2013 were 23 MtC whereas Wellington and Christchurch each emitted 0.4 MtC of fossil fuel CO_2 in 2013 (Boden et al., 2017; AECOM, 2016; Australian Government, 2016).

Deleted: figure 5

Deleted: 2012

Although broad-scale flow from the west is common, the local topography means that local air flow is almost always either southerly or northerly (Stephens et al., 2013), but during rare (<5% of the time) westerly wind events, fossil fuel emissions from Wellington do appear to cause enhancements of up to 2 ppm in CO_2 (Stephens et al., 2013), which would decrease $\Delta^{14}CO_2$ by ~1 ‰ during such an event. Yet there is no evidence of seasonality in the infrequent westerly events. Northerly conditions bring a terrestrial biosphere influence that elevates CO_2 by about 1 ppm (Stephens et al., 2013), which could result in a maximum increase in $\Delta^{14}CO_2$ of ~0.2‰ relative to background conditions, but there is no evidence that this influence is seasonally variable either. Thus, although there are some local influences on the Baring Head $\Delta^{14}CO_2$, none of these appear to be seasonally dependent and instead, the observed Baring Head $\Delta^{14}CO_2$ maximum in spring in the recent part of the record may be explained by the seasonal maximum in cross-tropopause exchange bringing ^{14}C-enriched air at this time of year.

Deleted: (figure 5)

5.3. Comparison with other atmospheric $\Delta^{14}CO_2$ records

Deleted: 4

We compare the Wellington $\Delta^{14}CO_2$ record with several other $\Delta^{14}CO_2$ records, located as indicated in figure 1. First, we compare with measurements from Cape Grim, Australia (CGO, 40.68°S, 144.68°E, 94 m asl). Cape Grim is at similar latitude to Wellington and also frequently receives air from the Southern Ocean (Levin et al., 2010). Samples are collected by a similar method to the Wellington record using NaOH absorption and are measured by gas counting to ~2 ‰ precision. Next, we compare with mid-latitude high-altitude clean air sites in the Northern Hemisphere. The Vermunt, Austria (VER, 47.07°N, 9.57°E, 1800 m asl) record began in 1958, only a few years after the Wellington record began, and in the 1980s the site was moved to Jungfraujoch, Switzerland (JFJ, 46.55°N, 7.98°E, 3450 m asl); these measurements are made in the same manner and by the same laboratory as the Cape Grim record (Levin et al., 2013). We also consider the Niwot Ridge, USA $\Delta^{14}CO_2$ record (NWR, 40.05°N, 105.59°W, 3523 m asl), which began in 2003 (Turnbull et al., 2007; Lehman et al., 2013). Niwot Ridge is also a mid-latitude high-altitude site, but samples are collected as whole air in flasks and measured by AMS in a similar manner to that described for the Wellington flask samples. Thus, we are comparing two independent Southern Hemisphere records with two independent Northern Hemisphere records, with the two hemispheres tied together by the common measurement laboratory used for Cape Grim and Jungfraujoch. Results from all records are compared in figure 5.

Deleted: 6

The Wellington and Cape Grim records are generally consistent with one another (Figure 3), with the exception of the 1995-2005 period, when the Wellington record is slightly higher, apparently due to bias in the Wellington record (discussed in section 3.5.3.). Differences between the sites are smaller than the measurement uncertainty for all other periods (table 2). This implies that $\Delta^{14}CO_2$ is homogeneous across Southern Hemisphere clean air sites within the same latitude band, at least since the 1980s when the two records overlap. Similarly, the high altitude, mid-latitude Northern Hemisphere sites are consistent with one another, although there are some differences in seasonal cycles in recent years (Turnbull et al., 2009b).

The bomb spike maximum is higher and earlier in the Northern Hemisphere records (figure 5), consistent with the production of most bomb ^{14}C in the Northern Hemisphere stratosphere. We make a new, simple estimate of the interhemispheric exchange time during the 1963 – 1965 period using the difference in the timing of the Northern and Southern Hemisphere bomb peaks. The first maximum of the bomb peak was in July 1963 in the Northern Hemisphere and January 1965 in the Southern Hemisphere, a 1.4 year offset, implying a 1.4 year exchange time. This is consistent with other more detailed interhemispheric exchange time estimates that have been determined from long-term measurements of SF_6 of 1.3 to 1.4 years (Geller at el., 1997; Patra et al., 2011).

Deleted: 6

Northern Hemisphere $\Delta^{14}CO_2$ remains higher than Southern Hemisphere $\Delta^{14}CO_2$ by about 20 ‰ until 1972. Although most nuclear weapons testing ceased in 1963, a few smaller tests continued in the late 1960s, contributing to this continued interhemispheric offset (Enting, 1982). The interhemispheric gradient disappeared within about 1.5 years after atmospheric testing essentially stopped in 1970. Except periods of noisy data from Vermunt in the late 1970s and Wellington in 1995-2005, there are only small (<2 ‰) interhemispheric gradients from 1972 until 2002 (figure 5, table 2).

Deleted: 6

As previously noted by Levin et al. (2010) using a shorter dataset, an interhemispheric gradient of 5-7 ‰ develops in 2002, with the Southern Hemisphere sites higher than the Northern Hemisphere sites (table 2). We choose 1986 – 1990 and 2005 – 2013 as time periods to compare, to avoid the periods where the Wellington record is noisy (1995 –

2005) and where we substituted flask measurements from 1990 – 1993. In 1986 – 1990, there is less than 2 ‰ difference between Wellington and either Cape Grim or Jungfraujoch. There is also no difference between the Cape Grim and Jungfraujoch records during this time period. The Wellington and Cape Grim records still agree within 2 ‰ after 2005, but both Jungfraujoch and Niwot Ridge diverge from Wellington, by 4.8 ± 2.7 and 6.9 ± 2.5 ‰, respectively; Jungfraujoch and Niwot Ridge are not significantly different from one another. This new interhemispheric gradient is robust, being consistent amongst the sites measured by three different research groups each with their own methods. It is not an artifact of interlaboratory offsets, since Cape Grim and Jungfraujoch measurements are made by the same group using the same sampling and measurement methods, and the Wellington and Niwot Ridge measurements (measured by different techniques) agree well with the other sites at similar latitude (Cape Grim and Jungfraujoch respectively). This developing gradient is also apparent in the larger sampling network of Levin et al (2010) and in a separate $\Delta^{14}CO_2$ sampling network (Graven et al., 2012), although that dataset extends only to 2007.

Graven et al. (2012) demonstrated that increasing (mostly Northern Hemisphere) fossil fuel CO_2 emissions cannot explain this $\Delta^{14}CO_2$ interhemispheric gradient, and instead, they postulated that ^{14}C uptake into the Southern Ocean reduced over time. Levin et al. (2010) were able to roughly replicate this interhemispheric gradient in their GRACE model by tuning the terrestrial biosphere fluxes to match the observed global average atmospheric CO_2 and $\Delta^{14}CO_2$ records. Where the observations suggest the rapid development of an interhemispheric gradient in the early 2000's (figure 5), the GRACE model simulates a more gradual transition over a period of roughly two decades. Independent evidence suggests that the Southern Ocean is more likely to be responsible for this rapid shift in the atmospheric $\Delta^{14}CO_2$ gradient. That is, an apparent reorganization of Southern Ocean carbon exchange in the early 2000s (Landschützer et al., 2015) is postulated to be associated with changes in upwelling of deep water (DeVries et al., 2017), to which atmospheric $\Delta^{14}CO_2$ is highly sensitive (Rodgers et al., 2011; Graven et al., 2012b). The observed $\Delta^{14}CO_2$ interhemispheric gradient is consistent with these postulated changes in upwelling. Other possible explanations for this new interhemispheric $\Delta^{14}CO_2$ gradient are substantial underreporting of Northern Hemisphere fossil CO_2 emissions (e.g. Francey et al., 2013) or changes in the land carbon sink (Wang et al., 2013; Sitch et al., 2015), although this latter is less likely since $\Delta^{14}CO_2$ is much less sensitive to biospheric fluxes than to either ocean or fossil fuel fluxes (e.g. Levin et al., 2010; Turnbull et al., 2009). Given the limited spatial coverage of the current $\Delta^{14}CO_2$ observing network, it is not possible to robustly determine which of these processes causes the interhemispheric gradient. This could be achieved with more observations of the spatial and temporal variations of atmospheric $\Delta^{14}CO_2$.

6. Conclusions

Deleted: 5

The 60 year-long Wellington $\Delta^{14}CO_2$ record has been revised and extended to 2014. Most revisions were minor, but we particularly note that the earlier reported 1990-1993 measurements have been entirely replaced with new measurements. A second period form 1995-2005 has poorer data quality than the rest of the record, and may also be biased high by a few permil. These data have been revised substantially, and new measurements have been added to this period, but we were unable to definitively identify or correct for bias, so the data have been retained, albeit with caution. We further validated the record by comparison with tree ring samples collected from the Baring Head sampling location and from nearby Eastbourne, Wellington; both tree ring records show excellent agreement with the original record, and indicate that there are no other periods where the original measurements are problematic.

The Wellington $\Delta^{14}CO_2$ time series records the history of atmospheric nuclear weapons testing and the subsequent decline of $\Delta^{14}CO_2$ as the bomb ^{14}C moved throughout the carbon cycle, and ^{14}C-free fossil fuel emissions further decreased $\Delta^{14}CO_2$. The timing of the first appearance of the bomb-^{14}C peak at Wellington is consistent with other recent estimates of interhemispheric exchange time at 1.4 years.

The seasonal cycle at Wellington evolves through the record, apparently dominated by the seasonality of cross-tropopause transport, which drives a changing seasonal cycle through time. In the early post-bomb period, the seasonally variable movement of bomb ^{14}C from the Northern Stratosphere through the Northern Troposphere to the Southern Troposphere appears to be the dominant control on the seasonal cycle at Wellington. The seasonal cycle reversed in later years, possibly due to a change in sign of the terrestrial biosphere $\Delta^{14}C$ signal. In recent years, the seasonal cycle has an amplitude of only 2 ‰, with a maximum in the austral spring. Cape Grim exhibits a similar seasonal cycle magnitude, but appears to be very slightly influenced by a terrestrial/anthropogenic signal during the austral winter that is not apparent at Wellington.

During the 1980s and 1990s, $\Delta^{14}CO_2$ was similar at mid-latitude clean air sites in both hemispheres, but since the early 2000s, the Northern Hemisphere $\Delta^{14}CO_2$ has dropped below the Southern Hemisphere by 5-7 ‰. The control on this changing interhemispheric gradient cannot be robustly determined from the existing sparse $\Delta^{14}CO_2$ observations, but may be due to a change in Southern Ocean dynamics reducing upwelling of old, ^{14}C-poor deep waters, consistent with recent evidence for an increasing Southern Ocean carbon sink. Alternative explanations are an underestimate of Northern Hemisphere fossil fuel CO_2 emissions, or a changing land carbon sink. This implies that ongoing and expanded Southern Hemisphere $\Delta^{14}CO_2$ observations and modelling may provide a fundamental constraint on our understanding of Southern Ocean dynamics and exchange processes.

7. Acknowledgements

Deleted: 6

A 60 year-long record takes more than a handful of authors to produce. This work was possible only because of the amazing foresight and scientific understanding of Athol Rafter and Gordon Fergusson, who began this record in the 1950s. Their work was continued over the years by a number of people, including Hugh Melhuish, Martin Manning, Dave Lowe, Rodger Sparks, Charlie McGill, Max Burr and Graeme Lyon. This work was funded by the Government of New Zealand as GNS Science Global Change Through Time core funding and NIWA Greenhouse Gases, Emissions, and Carbon Cycle Science Programme core funding. The author(s) wish to acknowledge the contribution of New Zealand eScience Infrastructure (NeSI) to the results of this research.

New Zealand's national computer and analytics services and team are supported by the NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation and Employment (http://www.nesi.org.nz). We thank Dr Scott Lehman (University of Colorado) and Dr Ingeborg Levin (University of Heidelberg) for sharing their $\Delta^{14}CO_2$ datasets for comparison with the Wellington record.

8. Data availability

Deleted: 7

The datasets presented in this paper are included as supplementary material. The datasets (including updates as they are available) can be accessed through the World Data Centre for Greenhouse Gases (http://ds.data.jma.go.jp/gmd/wdcgg/) or directly through GNS Science (https://gns.cri.nz/Home/Products/Databases/Wellington-atmospheric-14CO2-record) or NIWA (ftp://ftp.niwa.co.nz/tropac/).

9. References

AECOM New Zealand Limited, 2016. Community greenhouse gas inventory for Wellington City and the Greater Wellington Region 2000-2015, Wellington.

Australian Government, 2016. State and territory greenhouse gas inventories 2014. Department of the Environment.

Baisden, W.T., Prior, C.A., Chambers, D., Canessa, S., Phillips, A., Bertrand, C., Zondervan, A., Turnbull, J.C., 2013. Radiocarbon sample preparation and data flow at Rafter: Accommodating enhanced throughput and precision. Nuclear Instruments and Methods B294, 194-198.

Bevington, P.R., Robinson, D.K., 2003. Data reduction and error analysis for the physical sciences, Third Edition. McGraw-Hill.

Boden, T.A., Marland, G., Andres, R.J., 2017. Global, Regional, and National Fossil-Fuel CO_2 Emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Bozhinova, D., Combe, M., Palstra, S.W.L., Meijer, H.A.J., Krol, M.C., Peters, W., 2013. The importance of crop growth modeling to interpret the $^{14}CO_2$ signature of annual plants. Global Biogeochemical Cycles 27, 792-803.

Brailsford, G.W., Stephens, B.B., Gomez, A.J., Riedel, K., Mikaloff Fletcher, S.E., Nichol, S.E., Manning, M.R., 2012. Long-term continuous atmospheric CO_2 measurements at Baring Head, New Zealand. Atmospheric Measurement Techniques 5, 3109-3117.

Broecker, W.S., Peng, T.-H., Ostlund, H., Stuiver, M., 1985. The distribution of bomb radiocarbon in the ocean. Journal of Geophysical Research C4, 6953-6970.

Caldeira, K., Rau, G.H., Duffy, P.B., 1998. Predicted net efflux of radiocarbon from the ocean and increase in atmospheric radiocarbon content. Geophysical Research Letters 25, 3811-3814.

Cleveland, R., Cleveland, W., McRae, J., Terpenning, I., 1990. STL: A seasonal-trend decomposition procedure based on Loess. Journal of Official Statistics 6, 3-33.

Currie, K.I., Brailsford, G., Nichol, S., Gomez, A., Sparks, R., Lassey, K.R., Riedel, K., 2011. Tropospheric $^{14}CO_2$ at Wellington, New Zealand: the world's longest record. Biogeochemistry 104, 5-22.

Davies, T., Cullen, M., Malcolm, A., Mawson, M., Staniforth, A., White, A., Wood, N., 2005. A new dynamical core for the Met Office's global and regional modelling of the atmosphere. Quarterly Journal of the Royal Meteorological Society 131, 1759-2005.

DeVries, T., Holzer, M., Primeau, F., 2017. Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning. Nature 542, 215-218.

Djuricin, S., Pataki, D.E., Xu, X., 2010. A comparison of tracer methods for quantifying CO_2 sources in an urban region. Journal of Geophysical Research 115.

Enting, I.G., 1982. Nuclear weapons data for use in carbon cycle modelling. CSIRO Division of Atmospheric Physics and Technology, Melbourne, Australia.

Ferretti, D.F., Lowe, D.C., Martin, R.H., Brailsford, G.W., 2000. A new gas chromatograph-isotope ratio mass spectrometry technique for high-precision, N_2O-free analysis of $\delta^{13}C$ and $\delta^{18}O$ in atmospheric CO_2 from small air samples. Journal of Geophysical Research Atmospheres 105, 6709-6718.

Deleted: 7

Deleted: 2012

Francey, R.J., Trudinger, C.M., van der Schoot, M., Law, R.M., Krummel, P.B., Langenfelds, R.L., Steele, L.P., Allison, C.E., Stavert, A.R., Andres, R.J., Rödenbeck, C., 2013. Atmospheric verification of anthropogenic CO_2 emission trends. Nature Climate Change 3, 520-524.

Geller, L.S., Elkins, J.W., Lobert, J.M., Clarke, A.D., Hurst, D.F., Butler, J.H., Myers, R.C., 1997. Tropospheric SF_6: Observed latitudinal distribution and trends, derived emissions and interhemispheric exchange time. Geophysical Research Letters 24, 675-678.

Graven, H.D., Gruber, N., 2011. Continental-scale enrichment of atmospheric $^{14}CO_2$ from the nuclear power industry: potential impact on the estimation of fossil fuel-derived CO_2. Atmospheric Chemistry and Physics 11, 12339-12349.

Graven, H.D., Guilderson, T.P., Keeling, R.F., 2012. Observations of radiocarbon in CO_2 at seven global sampling sites in the Scripps flask network: Analysis of spatial gradients and seasonal cycles. Journal of Geophysical Research 117.

Hogg, A.G., 2013. SHCAL13 Southern Hemisphere calibration, 0-50,000 years CAL BP. Radiocarbon.

Hua, Q., Barbetti, M., Jacobsen, G., Zoppi, U., Lawson, E., 2000. Bomb radiocarbon in annual tree rings from Thailand and Australia. Nuc. Inst. and Meth. in Physics Research B 172, 359-365.

Hua, Q., Barbetti, M., Rakowski, A.Z., 2013. Atmospheric radiocarbon for the period 1950-2010. Radiocarbon 55, 1-14.

Jones, A.R., Thomson, D., Hort, M., Devenish, B., 2007. The UK Met Office's next-generation atmospheric dispersion model, NAME III, in: Borrego, C., Norman, A.-L. (Eds.), Air Pollution Modeling and Its Application XVII. Springer.

Kanu, A., Comfort, L., Guilderson, T.P., Cameron-Smith, P.J., Bergmann, D.J., Atlas, E.L., Schauffler, S., Boering, K.A., 2015. Measurements and modelling of contemporary radiocarbon in the stratosphere. Geophysical Research Letters 43.

Keeling, C.D., Piper, S.C., Whorf, T.P., Keeling, R.F., 2011. Evolution of natural and anthropogenic fluxes of atmospheric CO2 from 1957 to 2003. Tellus B 63, 1-22.

Keeling, C.D., Whorf, T., 2005. Atmospheric CO_2 records from sites in the SIO air sampling network, Trends: A compendium of data of global change. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, Oak Ridge, Tenn., USA.

Key, R.M., 2004. A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP). Global Biogeochemical Cycles 18.

Kjellström, E., Feichter, J., Hoffman, G., 2000. Transport of SF_6 and $^{14}CO_2$ in the atmospheric general circulation model ECHAM4. Tellus 52B, 1-18.

Landschützer, P., Gruber, N., Haumann, F.A., Rödenbeck, C., Bakker, D.C.E., van Heuven, S., Hoppema, M., Metzl, N., Sweeney, C., Takahashi, T., Tilbrook, B., Wanninkhof, R., 2015. The reinvigoration of the Southern Ocean carbon sink. Science 349, 1221-1224.

Law, R.M., Steele, L.P., Krummel, P.B., Zahorowski, W., 2010. Synoptic variations in atmospheric CO_2 at Cape Grim: a model intercomparison. Tellus B 62, 810-820.

Le Quere, C., Rodenbeck, C., Buitenhuis, E.T., Conway, T.J., Langenfelds, R., Gomez, A., Labuschagne, C., Ramonet, M., Nakazawa, T., Metzl, N., Gillett, N., Heimann, M., 2007. Saturation of the Southern Ocean CO_2 Sink Due to Recent Climate Change. Science 316, 1735-1738.

Lehman, S.J., Miller, J.B., Wolak, C., Southon, J.R., Tans, P.P., Montzka, S.A., Sweeney, C., Andrews, A.E., LaFranchi, B.W., Guilderson, T.P., Turnbull, J.C., 2013. Allocation of terrestrial carbon sources using $^{14}CO_2$: Methods, measurement, and modelling. Radiocarbon 55, 1484-1495.

Levin, I., Kromer, B., Hammer, S., 2013. Atmospheric $\Delta^{14}CO_2$ trend in Western European background air from 2000 to 2012. Tellus B 65.

Levin, I., Kromer, B., Schmidt, M., Sartorius, H., 2003. A novel approach for independent budgeting of fossil fuel CO_2 over Europe by $^{14}CO_2$ observations. Geophysical Research Letters 30, 2194.

Levin, I., Kromer, B., Schoch-Fischer, H., Bruns, M., Munnich, M., Berdau, D., Vogel, J.C., Munnich, K.O., 1985. 25 years of tropospheric ^{14}C observations in central Europe. Radiocarbon 27, 1-19.

Levin, I., Naegler, T., Kromer, B., Diehl, M., Francey, R.J., Gomez-Pelaez, A.J., Steele, L.P., Wagenbach, D., Weller, R., Worthy, D.E., 2010. Observations and modelling of the global distribution and long-term trend of atmospheric $^{14}CO_2$. Tellus B 62, 26-46.

Lopez, M., Schmidt, M., Delmotte, M., Colomb, A., Gros, V., Janssen, C., Lehman, S.J., Mondelain, D., Perrussel, O., Ramonet, M., Xueref-Remy, I., Bousquet, P., 2013. CO, NOx and $^{13}CO_2$ as tracers for fossil fuel CO_2: results from a pilot study in Paris during winter 2010. Atmospheric Chemistry and Physics 13, 7343-7358.

Lowe, D.C., Judd, W., 1987. Graphite target preparation for radiocarbon dating by accelerator mass spectrometry. Nuclear Instruments and Methods in Physics Research B 28, 113-116.

Manning, M.R., Lowe, D.C., Melhuish, W.H., Sparks, R.J., Wallace, G., Brenninkmeijer, C.A.M., McGill, R.C., 1990. The use of radiocarbon measurements in atmospheric sciences. Radiocarbon 32, 37-58.

Meijer, H.A.J., Pertuisot, M.-H., van der Plicht, J., 2006. High accuracy ^{14}C measurements for atmospheric CO_2 samples by AMS. Radiocarbon 48, 355-372.

Meijer, H.A.J., Smid, H.M., Perez, E., Keizer, M.G., 1996. Isotopic characterization of anthropogenic CO_2 emissions using isotopic and radiocarbon analysis. Physical Chemistry of the Earth 21, 483-487.

Miller, J.B., Lehman, S.J., Montzka, S.A., Sweeney, C., Miller, B.R., Wolak, C., Dlugokencky, E.J., Southon, J.R., Turnbull, J.C., Tans, P.P., 2012. Linking emissions of fossil fuel CO_2 and other anthropogenic trace gases using atmospheric $^{14}CO_2$. Journal of Geophysical Research 117, D08302.

Munro, D.R., Lovenduski, N.S., Takahashi, T., Stephens, B.B., Newberger, T., Sweeney, C., 2015. Recent evidence for a strengthening CO2sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002-2015). Geophysical Research Letters, n/a-n/a.

Naegler, T., Ciais, P., Rodgers, K., Levin, I., 2006. Excess radiocarbon constraints on air-sea gas exchange and the uptake of CO_2 by the oceans. Geophysical Research Letters 33.

Naegler, T., Levin, I., 2009. Observation-based global biospheric excess radiocarbon inventory 1963–2005. Journal of Geophysical Research 114.

Norris, M.W., 2015. Reconstruction of historic fossil CO_2 emissions using radiocarbon measurements from tree rings, School of Geography, Environment and Earth Sciences. Victoria University of Wellington.

Nydal, R., Lövseth, K., 1983. Tracing bomb ^{14}C in the atmosphere 1962-1980. Journal of Geophysical Research 88, 3621-3642.

Oeschger, H., Siegenthaler, U., Schotterer, U., Gugelmann, A., 1975. A box diffusion model to study the carbon dioxide exchange in nature. Tellus XXVII, 168-192.

Otago Daily Times, 1957. Polar ice caps may melt with industrialisation, Otago Daily Times, 23/1/1957 ed, Dunedin, New Zealand, p. 1.

Patra, P.K., Houweling, S., Krol, M., Bousquet, P., Belikov, D., Bergmann, D., Bian, H., Cameron-Smith, P., Chipperfield, M.P., Corbin, K., Fortems-Cheiney, A., Fraser, A., Gloor, E., Hess, P., Ito, A., Kawa, S.R., Law, R.M., Loh, Z., Maksyutov, S., Meng, L., Palmer, P.I., Prinn, R.G., Rigby, M., Saito, R., Wilson, C., 2011. TransCom model simulations of CH_4 and related species: linking transport, surface flux and chemical loss with CH_4 variability in the troposphere and lower stratosphere. Atmospheric Chemistry and Physics 11, 12813-12837.

Pickers, P.A., Manning, A.C., 2015. Investigating bias in the application of curve fitting programs to atmospheric time series. Atmospheric Measurement Techniques 8, 1469-1489.

Rafter, T.A., 1955. ^{14}C variations in nature and the effect on radiocarbon dating. New Zealand Journal of Science and Technology B37.

Rafter, T.A., Fergusson, G., 1959. Atmospheric radiocarbon as a tracer in geophysical circulation problems, United Nations Peaceful Uses of Atomic Energy. Pergamon Press, London.

Randerson, J.T., Enting, I.G., Schuur, E.A.G., Caldeira, K., Fung, I.Y., 2002. Seasonal and latitudinal variability of troposphere $\Delta^{14}CO_2$: Post bomb contributions from fossil fuels, oceans, the stratosphere, and the terrestrial biosphere. Global Biogeochemical Cycles 16, 1112.

Reimer, P.J., Brown, T.A., Reimer, R.W., 2004. Discussion: Reporting and calibration of post-bomb ^{14}C data. Radiocarbon 46, 1299-1304.

Rodgers, K.B., Mikaloff-Fletcher, S.E., Bianchi, D., Beaulieu, C., Galbraith, E.D., Gnanadesikan, A., Hogg, A.G., Iudicone, D., Lintner, B.R., Naegler, T., Reimer, P.J., Sarmiento, J.L., Slater, R.D., 2011. Interhemispheric gradient of atmospheric radiocarbon reveals natural variability of Southern Ocean winds. Climate of the Past 7, 1123-1138.

Sitch, S., Friedlingstein, P., Gruber, N., Jones, S.D., Murray-Tortarolo, G., Ahlström, A., Doney, S.C., Graven, H., Heinze, C., Huntingford, C., Levis, S., Levy, P.E., Lomas, M., Poulter, B., Viovy, N., Zaehle, S., Zeng, N., Arneth, A., Bonan, G., Bopp, L., Canadell, J.G., Chevallier, F., Ciais, P., Ellis, R., Gloor, M., Peylin, P., Piao, S.L., Le Quéré, C., Smith, B., Zhu, Z., Myneni, R., 2015. Recent trends and drivers of regional sources and sinks of carbon dioxide. Biogeosciences 12, 653-679.

Steinkamp, K., Mikaloff Fletcher, S.E., Brailsford, G., Smale, D., Moore, S., Keller, E.D., Baisden, W.T., Mukai, H., Stephens, B.B., 2017. Atmospheric CO_2 observations and models suggest strong carbon uptake by forests in New Zealand. Atmospheric Chemistry and Physics 17, 47-76.

Stephens, B.B., Brailsford, G.W., Gomez, A.J., Riedel, K., Mikaloff Fletcher, S.E., Nichol, S., Manning, M., 2013. Analysis of a 39-year continuous atmospheric CO_2 record from Baring Head, New Zealand. Biogeosciences 10, 2683-2697.

Stuiver, M., Polach, H.A., 1977. Discussion: Reporting of ^{14}C data. Radiocarbon 19, 355-363.

Stuiver, M., Quay, P.D., 1981. Atmospheric 14C changes resulting from fossil fuel CO2 release and cosmic ray flux variability. Earth and Planetary Science Letters, 53, 349-362.

Suess, H.E., 1955. Radiocarbon concentration in modern wood. Science 122, 414-417.

Sweeney, C., Gloor, E., Jacobson, A.R., Key, R.M., McKinley, G., Sarmiento, J.L., Wanninkhof, R., 2007. Constraining global air-sea gas exchange for CO_2 with recent bomb ^{14}C measurements. Global Biogeochemical Cycles 21.

Tans, P.P., De Jong, A.F., Mook, W.G., 1979. Natural atmospheric ^{14}C variation and the Suess effect. Nature 280, 826-828.

Thoning, K.W., Tans, P.P., Komhyr, W.D., 1989. Atmospheric carbon dioxide at Mauna Loa Observatory 2. Analysis of the NOAA GMCC data, 1974-1985. Journal of Geophysical Research 94, 8549-8563.

Trumbore, S.E., 2000. Age of soil organic matter and soil respiration: Radiocarbon constraints on belowground C dynamics. Ecological Applications 10, 399-411.

Turnbull, J.C., 2006. Development of a high precision $^{14}CO_2$ measurement capability and application to carbon cycle studies, Geological Sciences. University of Colorado, Boulder, p. 132.

Turnbull, J.C., Miller, J.B., Lehman, S.J., Hurst, D.F., Peters, W., Tans, P.P., Southon, J.R., Montzka, S.A., Elkins, J.W., Mondeel, D.J., Romashkin, P.A., Elansky, N.F., Shkorokhod, A., 2009. Spatial distribution of $\Delta^{14}CO_2$ across Eurasia: Measurements from the TROICA-8 expedition. Atmospheric Chemistry and Physics 9, 175-187.

Turnbull, J.C., Rayner, P.J., Miller, J.B., Naegler, T., Ciais, P., Cozic, A., 2009. On the use of $^{14}CO_2$ as a tracer for fossil fuel CO_2: quantifying uncertainties using an atmospheric transport model. Journal of Geophysical Research 114, D22302.

Turnbull, J.C., Sweeney, C., Karion, A., Newberger, T., Lehman, S.J., Tans, P.P., Davis, K.J., Lauvaux, T., Miles, N.L., Richardson, S.J., Cambaliza, M.O., Shepson, P.B., Gurney, K., Patarasuk, R., Razlivanov, I., 2015. Toward quantification and source sector identification of fossil fuel CO2 emissions from an urban area: Results from the INFLUX experiment. Journal of Geophysical Research: Atmospheres.

Turnbull, J.C., Zondervan, A., Kaiser, J., Norris, M., Dahl, J., Baisden, W.T., Lehman, S.J., 2015. High-precision atmospheric $^{14}CO_2$ measurement at the Rafter Radiocarbon Laboratory. Radiocarbon 57, 377-388.

Wang, Y., Li, M., Shen, L., 2013. Accelerating carbon uptake in the Northern Hemisphere: evidence from the interhemispheric difference of atmospheric CO_2 concentrations. Tellus B 65.

Ziehn, T., 2014. Greenhouse gas network design using backward Lagrangian particle dispersion modelling – Part 1: Methodology and Australian test case. Atmospheric Chemistry and Physics 14.

Zondervan, A., Hauser, T., Kaiser, J., Kitchen, R., Turnbull, J.C., West, J.G., 2015. XCAMS: The compact ^{14}C accelerator mass spectrometer extended for 10Be and

Deleted: Steinkamp, K., Mikaloff Fletcher, S.E., Brailsford, G., Smale, D., Moore, S., Keller, E.D., Baisden, W.T., Mukai, H., Stephens, B.B., 2016. Atmospheric CO_2 observations and models suggest strong carbon uptake by forests in New Zealand. Atmospheric Chemistry and Physics Discussions, 1-55.

26Al at GNS Science, New Zealand. Nuclear Instruments and Methods B361, 25-33.

Zondervan, A., Sparks, R.J., 1996. Development plans for the AMS facility at the Institute of Geological and Nuclear Sciences, New Zealand. Radiocarbon 38, 133-134.

10. Tables

Sampling Date Range	Sample ID NZ/NZA	Site	collection method	Measurement method
1954-1986	0-7500	MAK	tray	Gas counting
1987-1994	7500-8400	BHD	tray	Gas counting
1995-2004	8400-30000	BHD	bottle	AMS ENTandem ^{13}C ^{14}C
2005-2009	30000-34000	BHD	bottle	AMS ENTandem ^{12}C ^{13}C ^{14}C
2010-2011	34000-50000	BHD	bottle	AMS XCAMS
2012-present	50000-	BHD	bottle	AMS XCAMS/RG20

Table 1. Wellington $^{14}CO_2$ measurement methods through time. Gas counting samples are identified by NZ numbers, AMS samples by NZA numbers. NZ and NZA numbers do not ovelap. Sites are Makara (MAK) and Baring Head (BHD). Collection and measurement methods are described in detail in the text.

Site difference	**Time period**	**$\Delta^{14}CO_2$ difference (‰)**
BHD-CGO	1986-1990	1.8 ± 2.5
BHD-CGO	2005-2013	1.3 ± 3.4
BHD-JFJ	1986-1990	0.8 ± 3.9
BHD-JFJ	2005-2013	4.8 ± 2.7
BHD-NWR	2005-2013	6.9 ± 2.5

Table 2. $\Delta^{14}CO_2$ gradients between sites, determined as the mean of the monthly differences for each time period. Errors are the standard deviation of the monthly differences.

11. Figures

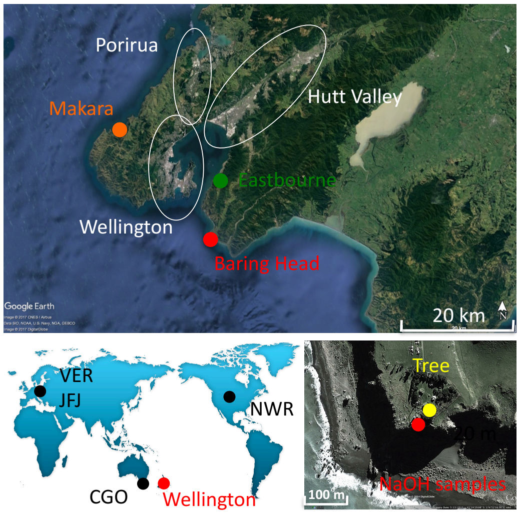

Figure 1. Sampling locations. Top: Makara (1954-1986) and Baring Head (1987 – present) air sampling sites, the location of the Eastbourne tree samples, and the urbanized areas of Wellington, Porirua and the Hutt Valley. Bottom left: world location showing Wellington and other sampling sites discussed in the text. Bottom right: close up of the Baring Head site showing the relative positions of the air (NaOH) and tree sampling locations.

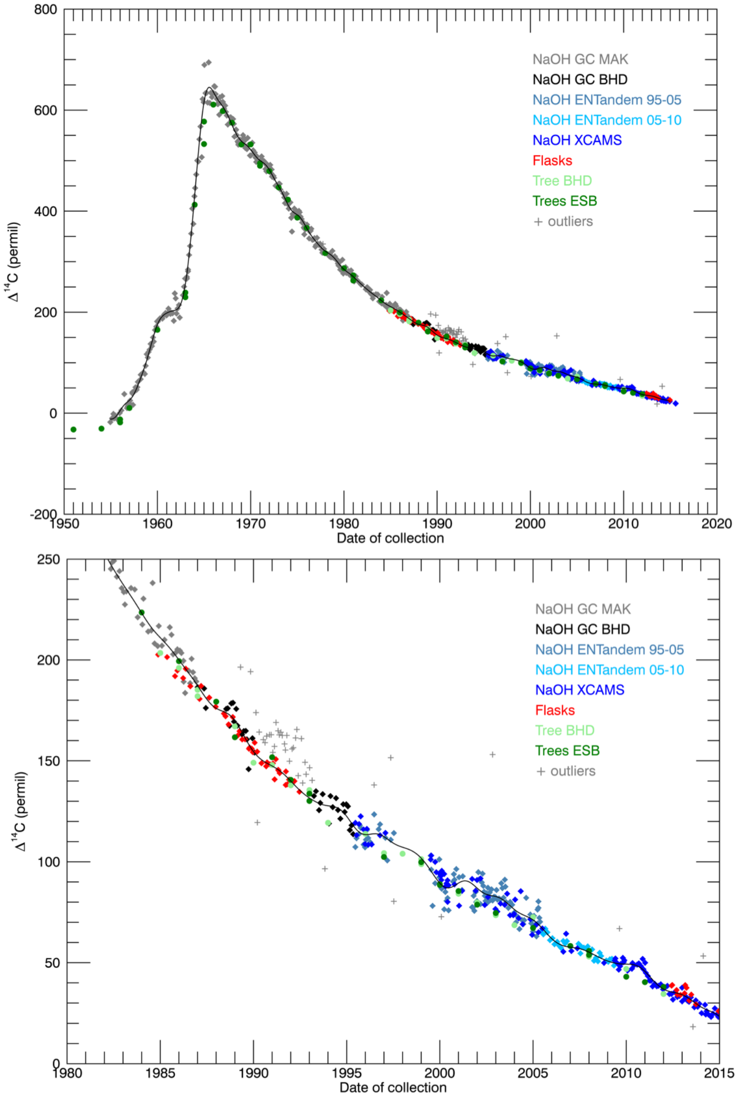

Figure 2. Wellington $^{14}CO_2$ record showing all collection and measurement methods for the full record (top) and zoomed in for the period since 1980 (bottom). Tree rings (green) and outliers (grey pluses) are excluded from the reported final dataset. Black line is the smooth curve fit to the final dataset.

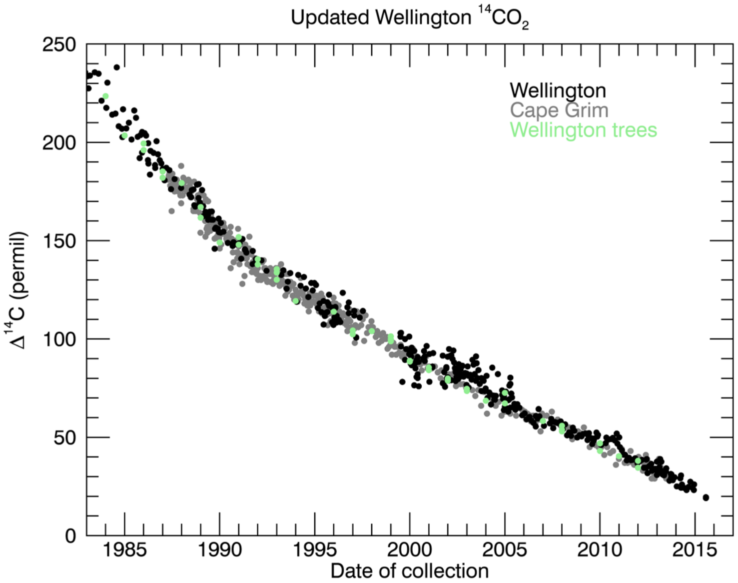

Figure 3. Comparison of the final Wellington and Cape Grim (Levin et al., 2010) $\Delta^{14}CO_2$ records. Wellington tree ring measurements are also shown.

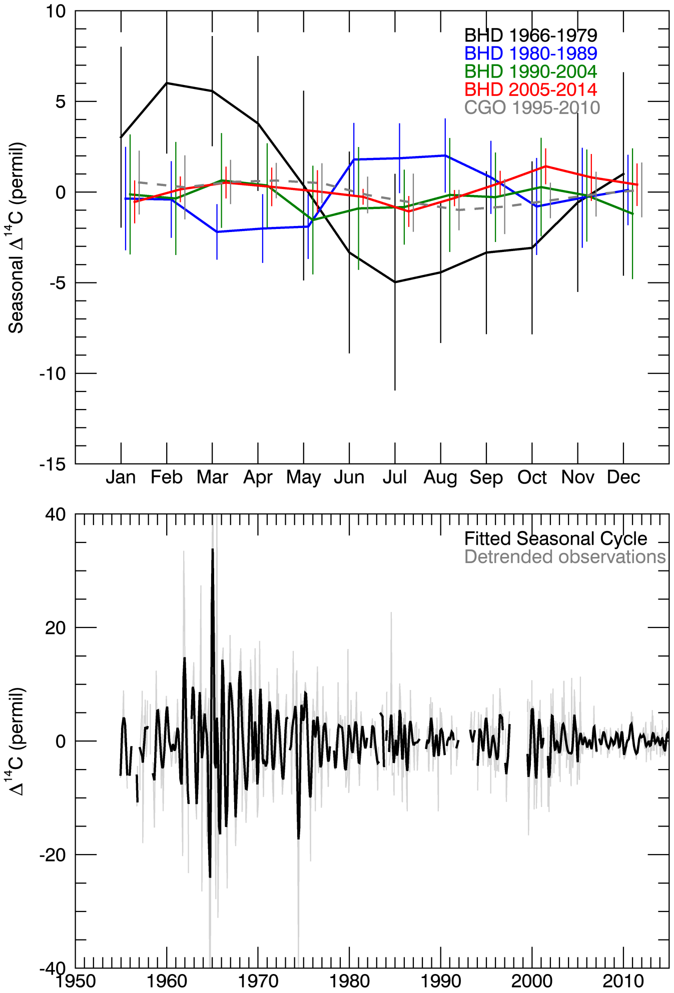

Figure 4. Detrended seasonal cycle in the Wellington $\Delta^{14}CO_2$ record. Top: BHD monthly detrended seasonal cycle averaged over four time periods as described in the text and the CGO (Levin et al., 2010) detrended seasonal cycle. Error bars are the standard deviation of all years averaged. Points for each time period are slightly offset for clarity. Bottom: full seasonal cycle record determined separately for each time period shown in the top panel plus 1954-1965 (black) and detrended observations without any smoothing (grey).

Deleted: Page Break

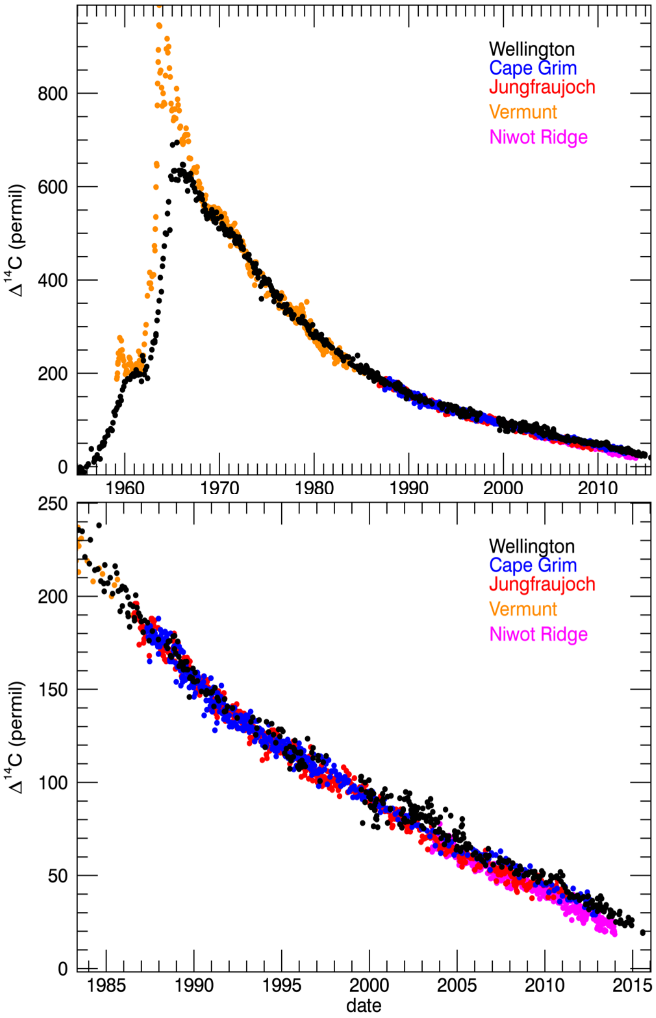

Figure 5. Comparison of Wellington and other atmospheric Δ^{14}CO$_2$ records (Levin et al., 2010; Turnbull et al., 2007; Lehman et al., 2013).

Deleted: 5

Supplementary material: Sixty years of radiocarbon dioxide measurements at Wellington, New Zealand 1954 – 2014

Wellington $^{14}CO_2$ record methods and updates

An overview of the measurement methods and summary table is given in the main paper (table 1). This supplementary material provides full details of the sampling methods used through time, compiling methodological information provided in previous reports on the Wellington record (Rafter and Fergusson 1959, Manning et al., 1990, Currie et al., 2011) along with methods newly applied in this new extension and refinement of the dataset. This supplementary material intentionally contains some repetition from the main paper so that all methodological details are available in a single document for users of the dataset.

1. Sampling sites

Samples from 1954 – 5 June 1987 were collected at Makara, on the west coast of the North Island (MAK, 41.25°S, 174.69°E, 300 m asl). The sampling site was moved to Baring Head on the South Coast of the North Island and 23 km southeast of Makara (BHD, 41.41°S, 174.87°E, 80 m asl). All samples since 8 July 1988 have been collected at Baring Head (figure 1). A set of tree ring samples was collected from two kauri trees from Nikau St, Eastbourne (NIK, 41.31°S, 174.89°E, 5 m asl). Sampling site for each sample is included in the dataset. The overall dataset uses site code BHD to be consistent with other greenhouse gas and isotope measurements from Baring Head.

2. Date of collection

Samples are collected over a period of several weeks, so the date of collection is determined as the middle of the sampling period. In some cases, the dates of start and end were not recorded or there is ambiguity as to whether the date recorded is the middle date, start date or end date.

For gas counting records, usually the date of collection was taken from Currie et al. (2011). These are usually the same as the dates reported in Manning and Melhuish (1990), but some samples reported at the end of the Manning and Melhuish (1990) dataset used the start date as date of collection (rather than the mean date), and this was corrected in Currie et al. (2011). For the present update, additional sample date information for a number of early measurements (1950s-1970s) was recorded in a comment field in the GNS Science radiocarbon database (RLIMS), and the date of collection was re-determined from this information.

For the AMS measurements starting in 1995, the date of collection was not included in RLIMS in most cases. We used Currie et al. (2011) and a separate sampling spreadsheet log kept at NIWA to determine the date of collection.

For samples where both start and end dates were recorded, the mean date is used as the date of collection. Where only the start date was recorded, we use two options. For samples measured by gas counting, we use the start date (as the number of days exposed was variable). For samples measured by AMS (dates after 1995), there are only a few for which the end date is missing:

- For NZA9240 (start date 1/4/96), NZA11523 (start date 15/6/96) and NZA11524 (start date 1/12/99). We assume a 14-day collection period since most other samples during this period were collected over 14 days.
- NZA14572 (start date 11/9/00) and NZA14573 (start date 22/9/00). We assume the end date is the same date as the start date for the following sample, which give days exposed of 11 and 18 days respectively.

3. Collection methods

3.1. Static NaOH absorption

The primary collection method is static absorption of CO_2 into sodium hydroxide (NaOH) solution, which is left exposed to air at the sampling site for ~2 weeks. The NaOH solution is first prepared to remove any CO_2 that was initially present when making up the solution. The standard preparation method involves bubbling CO_2-free air or nitrogen through deionized water for ~5 hours. We continue using this step, although it may be unnecessary since any initially-present CO_2 is removed in the following steps. NaOH pellets are then added to make 0.5 or 1 M. The NaOH pellets always absorb some ambient CO_2 when exposed to air, inevitably introducing some CO_2 contamination during the preparation, and this is removed from the solution by the addition of ~ 1 g barium chloride (0.004 mol) in ~50 mL of bubbled, deionized water. Any carbonate is precipitated as barium carbonate and the resultant nominally carbonate-free solution is decanted into smaller bottles under a stream of CO_2-free air or nitrogen. Since 2015, a further step adds 0.2 mL of 85% phosphoric acid to the solution after the barium chloride precipitation step. Any remaining barium is then precipitated as barium phosphate, ensuring that all the CO_2 absorbed during sample collection remains as carbonate rather than precipitating out as barium carbonate. The resulting solution is stored in sealed bottles until ready for use.

From 1954-1995, ~ 2 L of the 0.5 M NaOH solution was poured into a large (~450 cm^2 surface area) Pyrex® tray and left exposed to air for 1 - 2 weeks. Some early (1954-1970) samples may have been collected using different vessels. In a few early samples, air may have been pumped through the solution rather than passive absorption (Manning et al., 1990). Two samples were collected using $Ba(OH)_2$ rather than NaOH, in 1957 and 1958. As these are not outliers, they are included in the dataset, with a comment and flag.

Beginning in 1995, when the measurement technique changed from gas counting to AMS measurement (see section 3), the large trays were replaced with high density polyethylene (HDPE) bottles containing ~200 mL 0.5 M NaOH solution, with the depth of liquid kept the same as it was in the previously used trays. The lid is removed and the bottle is left open inside a Stevenson meteorological screen for ~2 weeks. Date of opening and closing bottle is recorded. The two methods (large tray and small bottle) have been compared, with no significant difference observed between the two methods (Currie et al., 2011).

Static NaOH absorption necessarily fractionates relative to CO_2 in the atmosphere; typical $\delta^{13}C$ values are -15 to -25 ‰ for these samples. This is corrected for in the data analysis.

3.2. Whole air flask samples

Flasks of whole air are collected by flushing ambient air through the flask for several minutes. The flask is then closed and allowed to fill to slightly over ambient pressure. All flask samples presented here were collected at the Baring Head site. Most of those collected during 1984 – 1993 were collected during southerly, clean air, conditions (Stephens et al., 2013). Flasks collected since 2013 were all collected during southerly conditions.

3.3. Tree rings

When trees photosynthesize, they faithfully record the $\Delta^{14}C$ of ambient CO_2 in their cellulose, the structural component of wood. Annual tree rings therefore provide a summertime (approximately September – April in the Southern Hemisphere) daytime average $\Delta^{14}CO_2$. Photosynthetic uptake varies during the daylight hours depending on factors including growth period, sunlight, and temperature (Bozhinova et al., 2013). For the Wellington location with typically high wind speeds and little short-term variability in CO_2 and other trace gases, this variability in photosynthesis is likely not a big impact on $\Delta^{14}CO_2$. The ring age we assign differs from standard nomenclature for Southern Hemisphere dendrochronology. We assign the mean age of the ring as January 1 of the year in which growth finished (i.e. the mean age of a ring growing from September – April), whereas dendrochronologists assign the "ring year" is as the year in which ring growth started (i.e. the previous year).

We collected cores from three trees close to the Baring Head site. A pine (*Pinus radiata*) located 10 m from the Baring Head sampling station (figure 1) yielded rings back to 1986 (Norris, 2015). New Zealand kauri (*Agathis australis*) is a long-lived hardwood species that has been widely used in dendrochronology and radiocarbon calibration studies (Hogg et al., 2013). The rings of this species are well defined and seldom missing, making this an ideal species for our comparison. We collected cores from two specimens planted in 1919 and 1920, located 20 m from one another in Eastbourne, 12 km from Baring Head (figure 1).

We evaluated the ring counts by the timing of the ^{14}C bomb spike. Shifting the Eastbourne record by one year in either direction moves the maximum in the tree ring record out of phase with the final Wellington $\Delta^{14}CO_2$ record (figure S1), confirming that the ring counts are correct. For the Baring Head pine, rings go back to only 1986, so we compare with the Eastbourne record, and find that the two records agree very well, with an insignificant mean difference of -0.4 ± 0.8 ‰ (figure S1).

Sampling of the tree rings for ^{14}C analysis is also a potential source of bias; ideally only the full ring for a given year would be sampled. In practice, it is difficult to ensure that the full ring is isolated without losing any material from that ring and no wood from the surrounding rings is included. To evaluate this, we measured replicate samples different cores from the same tree (Baring Head pine) or two different trees (Eastbourne kauri). For samples collected since 1985, all these replicates are consistent within their uncertainties (figure S2). We also tested shifting the Baring Head pine ring counts by one year in either direction, and find that when we do so, the mean difference between the Baring Head pine and Eastbourne kauri is +5.6 ± 0.7 ‰ or -8.4 ± 0.8 ‰ when moving the chronology one year back or forward respectively. These differences are too large to be plausible differences between these two sites, and give us further confidence that the ring counts are correct.

We do find large differences of 9.2, 44.5 and 4.9 ‰ for three replicates from Eastbourne in 1963, 1965 and 1971, which we attribute to small differences in sampling of the rings which were magnified by the rapid change in $\Delta^{14}C$ of up to 200 ‰ yr^{-1} during this period. Thus, the ring values during this period should be treated with caution.

Deleted: ‰ for

4. ^{14}C Analysis methods

4.1. Extraction

From 1954 - 1995 CO_2 was extracted from the NaOH solution by acidification followed by cryogenic distillation (Rafter and Fergusson, 1959). AMS sample extraction from 1995 onward follows the general same method, but with smaller volumes (Currie et al., 2011), and this extraction is performed at NIWA. The $\delta^{13}C$ is measured at NIWA on an aliquot of CO_2 from the same extraction.

CO_2 is extracted from whole air flask samples by cryogenic extraction. Samples collected from 1984- 1993 were extracted at GNS Science by Graeme Lyon and archived as ampoules of pure CO_2. In 2012, these tubes were cracked under vacuum to liberate the CO_2. Any leakage during storage is readily by air present in the tube when it is cracked for transfer and tubes with leakage were discarded. An aliquot of each was measured for $\delta^{13}C$ to confirm that no fractionation had occurred during storage. Whole air samples collected since 2013 are analyzed for $\delta^{13}C$ and other trace gases and isotopes at NIWA and for the $^{14}CO_2$ measurement, CO_2 is extracted from whole air at GNS Science (Turnbull et al. 2015).

Cellulose was isolated from whole tree rings by first removing labile organics with a series of solvent washes (hexane, isopropanol and acetone) in a Soxhlet system (Norris, 2015). This was followed by oxidation to isolate the cellulose from other materials (Hua et al., 2000). The cellulose was combusted and the CO_2 purified following standard methods in the GNS Science laboratory (Baisden et al., 2013).

4.2. Graphitization (AMS only)

The LG1 graphitization system was used from 1995 to 2011 (NZA < 50,000) (Lowe et al., 1987), and replaced with the RG20 graphite system in 2011 (NZA > 50,000) (Turnbull et al., 2015). In both systems, CO_2 is reduced to graphite over iron catalyst in the presence of hydrogen gas. RG20 has online pressure monitoring to ensure the reaction goes to completion. On LG1, iron powder was used as catalyst; on RG20 iron oxide is reduced to iron online immediately prior to graphitization. All samples measured on the EN

Tandem were prepared using LG1, and all except one sample measured on XCAMS were prepared using RG20.

4.3. Decay counting

Static NaOH samples were measured by conventional decay counting on the CO_2 gas from 1954-1995, described in detail in earlier publications (Manning and Melhuish, 1990; Currie et al., 2011). Samples measured by gas counting are identified by unique numbers with the prefix “NZ”.

4.4. AMS

All measurements made since 1995, including recent measurements of flask samples collected in the 1980s and 1990s, were measured by accelerator mass spectrometry (AMS). The unique identifiers for these samples have the prefix “NZA”, but the NZ and NZA numbers never overlap.

From 1995 to 2010, an EN Tandem AMS was used for measurement (NZA < 35,000, Zondervan and Sparks, 1996). Four graphite targets were prepared from each sample by splitting a single large CO_2 aliquot under equilibrium conditions, then reducing to graphite using the LG1 graphite system. Until 2005 (NZA <30,000), only ^{13}C and ^{14}C were measured on the EN Tandem system, so the normalization correction for isotopic fractionation (Stuiver and Polach, 1977) was performed using the IRMS $\delta^{13}C$ value of the sample measured at NIWA. The reduction in scatter from 2005 onwards (NZA > 30,000) reflects the addition of online ^{12}C measurement in the EN Tandem system. This allows direct online correction for isotopic fractionation that may occur during sample preparation and in the accelerator itself (Zondervan et al., 2015), and results in improved long-term repeatability.

Measurements of multiple targets are averaged to provide a weighted mean $F^{14}C$ and $\Delta^{14}C$. This averaging was done outside of the AMS database, and records of how these original calculations were done exist for only some of the results. In many cases, one or two targets appeared to be outliers and were excluded from the averaging calculation. In some cases, simple means of the multiple targets were used instead of weighted means.

We have recalculated the results for all of these multiple target samples, using consistent criteria for all. To determine and flag outliers, the reduced chi square statistic (χ^2_v) was calculated for each individual sample date (typically four targets) and for the full multi-target dataset. Note that in all except three cases, the multiple targets were all measured within the same AMS wheel, so wheel-to-wheel variability is not accounted for here.

First, including all multi-target measurements, we obtain χ^2_v =1.2 from 598 targets. We identified outliers as those where χ^2_v for that sample date was greater than three and excluded the targets within that set for which the residual was greater than two. Recalculating χ^2_v after removing these outliers obtains χ^2_v=0.99 from 585 targets, a satisfactory result. From EN Tandem measurements only, χ^2_v=1.02 from 529 targets. For XCAMS, multiple targets were collected from both within and between wheels. The Baring Head NaOH samples measured on XCAMS give, χ^2_v=0.95 from 56 targets. The weighted means have been recalculated from the $F^{14}C$ values, and $\Delta^{14}C$ calculated from $F^{14}C$.

For samples where the outliers changed, differences of up to 5‰ from the values reported by Currie et al. (2011) occur. For other samples, differences of 0.1 – 0.5 ‰ occur, apparently because when results were initially calculated, straight rather than weighted means (weighted by reported uncertainties) were used in some cases and rounding errors in other cases.

A few samples were measured on the EN Tandem in two different wheels, and in some cases were assigned different R (laboratory) numbers. In those cases, Currie et al. (2011) did not always average all four replicates. This has now been revised.

In 2010, the EN Tandem was replaced with a National Electrostatics Corporation AMS, dubbed XCAMS (NZA > 34,000). XCAMS measures all three carbon isotopes, such that the normalization correction is performed using the AMS measured ^{13}C values (Zondervan et al., 2015). XCAMS measurements are made on single graphite targets, measured to high precision, typically better than 2‰ overall uncertainty. Replicate analyses within and between wheels have been used to determine long-term repeatability and overall uncertainty levels (Turnbull et al., 2015), including replicate measurements of samples from this dataset. Where more than one measurement was made for a given date, we report the weighted mean of all measurements.

Three different preparation methods have been used for the primary standards (oxalic acid I, OxI) using in determining $\Delta^{14}CO_2$. Sealed tube combustion was used for all the EN Tandem measurements, whereby sufficient solid OxI material was combusted to provide five or six aliquots. The resultant CO_2 was split and then graphitization was performed separately for each aliquot. Since 2010, OxI has also been prepared by combustion of single aliquots by elemental analyzer and subsequent cryogenic CO_2 collection (Baisden et al., 2013). Both these methods show larger scatter than air standard materials, and there is an offset of -1.5‰ in samples measured with OxI made by sealed tube combustion, and those samples have been corrected for this offset (Turnbull et al., 2015). To reduce the scatter in OxI due to variability in combustion (which air samples do not undergo), we prepared two large flasks of CO_2 from OxI, and in 2015, began taking aliquots of this CO_2 as primary standard material. There is no offset relative to the elemental analyzer combustion method, and repeatability is slightly improved.

During the transition between the EN Tandem and XCAMS instruments, three Wellington $^{14}CO_2$ samples were measured on both instruments, and XCAMS results were 4.2‰ higher for these three samples. However, this jump during the transition does not appear consistent across wheels, and the larger quality control dataset which shows no significant difference in several solid sample quality control standards (Baisden et al., 2013), and is not addressed further here.

5. Results format

5.1. Date of collection

NaOH samples are collected over a period of typically two weeks, and sometimes much longer. We determine the date of collection as the average of the start and end collection dates. In cases where the end date was not recorded, we use the start date. For a few samples, the sampling dates were not recorded or are ambiguous, and those results have been flagged and excluded from the final dataset.

5.2 $\Delta^{14}C$ and $F^{14}C$

Results are reported here in $F^{14}C$ (Reimer et al., 2004) and $\Delta^{14}C$ (Stuiver and Polach, 1977). $F^{14}C$ is corrected for isotopic fractionation and blank corrected. $F^{14}C$ is calculated from the original measurement data recorded in our databases, and has been updated for only a handful of records where errors were found in the $\delta^{13}C$ value originally used in the normalization correction. No changes in measured values were made except in cases where multiple targets were averaged.

$\Delta^{14}C$ is derived from $F^{14}C$, with a correction for radioactive decay since the time of collection (Stuiver and Polach, 1977). The date of collection is required for this calculation. In previously published versions of the dataset, the gas counting and EN Tandem results used only the year of collection in the $\Delta^{14}C$ calculation, and in many cases, $\Delta^{14}C$ was initially reported using the date of measurement. Currie et al. (2011) re-calculated $\Delta^{14}C$ in a separate database, using the date of collection, but assumed that the date of measurement, rather than the year of measurement, had been used. Here we went back to the original measured $F^{14}C$, and to the sample collection database to find the date of collection. All $\Delta^{14}C$ results are recalculated using the decimal date of collection. This results in changes of a few tenths of permil in most $\Delta^{14}C$ values relative to those reported by Currie et al (2011) and Manning and Melhuish (1990). Samples for which changes have been made relative to the previously published results are indicated by the third column quality flag provided in the supplementary dataset.

5.3. Uncertainties

The primary source of uncertainty is the ^{14}C counting statistical uncertainty determined from the number of beta decays (gas counting) or ^{14}C counts (AMS). For AMS measurements, we add an additional error term, determined from the long-term repeatability of secondary standard materials and added in quadrature to the AMS uncertainty. In early AMS measurements, this was 4%, and has decreased through time to 0.12%. For the EN Tandem results, this was determined from the performance of OxI primary standard targets through time, which may somewhat overestimate the uncertainty for air samples (Turnbull et al., 2015). For XCAMS measurements, long-term repeatability of air samples has been assessed based on repeated splits of CO_2 from a subset of these NaOH samples, as well as repeated extractions of aliquots of CO_2 from pressurized tanks of whole air. These repeated aliquots are measured both within the same AMS wheel and across multiple wheels. Within individual wheels, the counting statistical uncertainty is sufficient to explain the full variability, but an additional uncertainty of 0.12% is required to explain the spread across multiple wheels (Turnbull et al., 2015).

6. Smooth curve fits

Formatted: Heading 2

In addition to the raw measured $\Delta^{14}CO_2$ values, we calculate a smooth curve fit ("fit" in the supplementary datasets) and deseasonalized trend from the Wellington $\Delta^{14}C$ and $F^{14}C$ datasets. The deseasonalized trend ("trend" in the supplementary datasets) may be more useful than the raw data for aging of recent materials (e.g. Reimer et al., 2004; Hua et al., 2013). Acknowledging that the 1995-2005 period is variable and possibly biased in the Wellington record, we also provide in the supplementary material an alternative mid-latitude Southern Hemisphere smooth curve fit and deseasonalized trend in which the Wellington data for 1995-2005 has been removed and replaced with the Cape Grim data for that period (Levin et al., 2010).

Curvefitting is particularly challenging for the $\Delta^{14}CO_2$ record, since (a) there are data gaps and inconsistent sampling frequency, (b) the growth rate and trend vary dramatically and (c) the seasonal cycle changes both in magnitude and phase (section 4.2). We chose to use the CCGCRV fitting procedure (Thoning et al., 1989; www.esrl.noaa.gov/gmd/ccgg/mbl/crvfit/), which uses fast Fourier transform and low-pass filtering techniques to obtain a smoothed seasonal cycle and long term trend from atmospheric data. This technique can readily handle the data gaps, inconsistent sampling frequency, and rapid changes in the seasonal cycle and trend.

CCGCRV assigns a single set of harmonic terms across the full time period, which is inappropriate in this case of large variation in the seasonal cycle. Thus, we separate the record into five time periods: 1954-1965, 1966-1979, 1980-1989, 1990-2004, 2005-2014. These divisions were chosen based on major changes in the raw observational growth rate, seasonal cycle and data quality. The peak of the bomb spike in the Southern Hemisphere (1965) results in a very large change in seasonality that makes an obvious cutoff point. There is an obvious change in seasonality in the raw observations in 1979 – 1980. The 1990 to 2004 period was grouped to include the time when flask measurements have supplemented original NaOH measurements, and the 1995 – 2004 period with noisy data.

The other widely used fitting procedure, seasonal trend decomposition using locally weighted scatter plot smoothing (STL, Cleveland et al. 1990; Pickers et al., 2015) assumes that the seasonal cycle and trend change only gradually over a specified time period. This assumption is problematic for the $\Delta^{14}CO_2$ time-series, due to the rapid changes in the trend and seasonal amplitude during and following the bomb spike. Using this method would necessitate both gap-filling the record and dividing the record into time periods (as we have done for CCGCRV), giving no advantage over CCGCRV.

For each time period, we use CCGCRV with one linear and two harmonic terms and fit residuals are added back using a low-pass filter with an 80-day cutoff in the frequency domain. At each transition, we overlapped a two-year period and linearly interpolated the two fits across that two-year period to smooth the transitions caused by end effects. We tested different overlap periods, and found that two years was optimal to minimize end effects and retain the benefit of separating the time periods. The deseasonalized trend was determined from the full dataset rather than the five time periods, as it does not include the seasonality and produces the same result in either case.

We tested other time period divisions, and our chosen time divisions have the lowest mean residual difference from the measured $\Delta^{14}\mathrm{CO}_2$, indicating the best fit to the data (we tested only periods of >10 years since it is difficult to draw conclusions about seasonal cycles from shorter periods when the seasonal cycle amplitude is small relative to the measurement uncertainty). The mean difference between the fitted curve and the measured $\Delta^{14}\mathrm{CO}_2$ values is 3.8 ‰, consistent with the typical measurement uncertainty for the full dataset. Further, the residuals are highest for the early period (1954-1970) at 6 ‰, consistent with the larger measurement errors at that time of ~6 ‰. The residuals improve as the measurement errors reduce, such that since 2005, the mean residual is 2 ‰, consistent with the reported 2 ‰ uncertainties. The exception is the 1995- 2005 period where the mean residual difference of 5 ‰ is substantially higher than the mean reported uncertainty of 2.5 ‰, reflecting the apparent larger scatter during this period as discussed in section 3.5.3.

The one-sigma uncertainty on the smoothed curve and deseasonalized trend were determined using a Monte Carlo technique (n=100). Each data point was perturbed by a random normal error based on the reported uncertainty of that data point, such that the standard deviation of all perturbations would equal the reported uncertainty to derive the one-sigma uncertainty for the smooth curve. This is provided for further users of the dataset, and may be particularly helpful when the dataset is used for aging of recent materials.

Deleted: .

Deleted: 6

6. Data validation

6.1. Tree ring comparison

Over the more than 60 years of the Wellington $\Delta^{14}\mathrm{CO}_2$ record, there have necessarily been many changes in methodology, and the tree rings provide a way to validate the full record, albeit with lower resolution. Due to the possible sampling biases in the tree rings, use them to validate the existing measurements, but do not include them in the main Wellington $\Delta^{14}\mathrm{CO}_2$ dataset. We include the tree ring data in a separate table for researchers who may want to use these data in addition to the main Wellington $\Delta^{14}\mathrm{CO}_2$ dataset.

During the rapid $\Delta^{14}\mathrm{CO}_2$ change in the early 1960s, there are some differences between the kauri tree ring and Wellington $\Delta^{14}\mathrm{CO}_2$ records. The 1963 and 1964 tree ring samples are slightly lower than the concurrent $\Delta^{14}\mathrm{CO}_2$ samples. The peak $\Delta^{14}\mathrm{CO}_2$ measurement in the tree rings is 30‰ lower than the smoothed $\Delta^{14}\mathrm{CO}_2$ record, and 100‰ lower than the two highest $\Delta^{14}\mathrm{CO}_2$ measurements in 1965. These differences are likely due to small errors in sampling of the rings, which will be most apparent during periods of rapid change.

Prior to 1960 and from the peak of the bomb spike in 1965 until 1990, there is remarkable agreement between the tree rings and Wellington $\Delta^{14}\mathrm{CO}_2$ record, with the wiggles in the record replicated in both records. And since 2005, there is excellent agreement across all the different records. Some differences are observed in 1990-1993 and 1995-2005, which we discuss in the following sections.

6.2. 1990-1993 excursion

An excursion in the gas counting measurements between 1990 and 1993 has previously been noted as a deviation from the Cape Grim $\Delta^{14}CO_2$ record (Levin et al., 2010) during the same period. Cape Grim is at a similar latitude and also draws air from Southern Ocean. The Wellington and Cape Grim records overlap during almost all other periods.

We use archived CO_2 from flask samples to evaluate this period of deviation. First, the recent flask samples agree very well with the NaOH static samples from the same period (figure 2). This indicates that despite the difference in sampling period for the two methods, flask samples reflect the $\Delta^{14}CO_2$ observed in the longer-term NaOH static samples. We selected a subset of the archived 1984 - 1992 CO_2 samples for measurement, mostly from Southerly wind conditions, but including a few from other wind conditions. Each sample was liberated and the $\delta^{13}C$ measured before $\Delta^{14}CO_2$ measurement was performed. A few of the samples had been compromised, and this was readily apparent as they contained air rather than CO_2. $\delta^{13}C$ from all the remaining samples was in agreement with $d^{13}C$ measured from separate flasks collected at Baring Head and measured for $\delta^{13}C$ by Scripps at close to the time of collection. These flask $\Delta^{14}CO_2$ measurements do not exhibit the excursion seen in the NaOH static samples, implying that the deviation observed in the original NaOH static samples may be a sampling error. Annual tree rings from both the kauri and pine follow the flask measurements for this period (figure 2), confirming that the NaOH static samples are anomalous.

The 1990-1993 period was characterized by major changes in New Zealand science, both in the organizational structure and personnel. Although we are unable to exactly reconstruct events that time, we hypothesize that the NaOH solution preparation was conducted slightly differently, perhaps omitting the $BaCl_2$ precipitation step for these samples. This would result in contaminating CO_2 absorbed on the NaOH before the solution was prepared, which would result in higher $\Delta^{14}CO_2$ observed in these samples than in the ambient air. In any case, these values are anomalous and we remove the original NaOH static sample measurements between 1990 and 1993 and replace them with the new flask measurements for the same period.

6.3. 1995-2005 variability

In 1995, the measurement method was changed from gas counting to AMS. During the first ten years of AMS measurements, the record is much noisier than during any other period (figure 2). In 2005, online ^{12}C measurement was added to the AMS system, substantially improving the measurement accuracy, and the noise in the $\Delta^{14}CO_2$ record immediately reduced.

The remaining NaOH solution for all samples collected since 1995 has been archived, and typically only every second sample collected was measured, with the remainder archived without sampling. In 2011-2016, we revisited the 1995-2005 period, remeasuring some samples that had previously been measured and some that had never been measured for a total of 52 new analyses.

The new measurements on this period do show reduced scatter over the original analyses, particularly for the period from 1998-2001 where the original analyses appear anomalously low and in 2002-2003 when the original analyses appear anomalously high. Yet there remain a number of both low and high outliers in the new measurements. These are present in both the samples that were remeasured and in those for which this was the first sample from the bottle. This suggests that a subset of the archived sample bottles were either contaminated at the time of collection, or that some bottles were insufficiently sealed, causing contamination with more recent CO_2 during storage. Comparison with the tree ring measurements and with the Cape Grim record (Levin et al., 2010) suggest that the measurements during this period may, on average, be biased high as well as having additional scatter (figure 3). Nonetheless, in the absence of better data, we retain both the original and remeasured NaOH sample results in the full record.

8. Existing results added to the Wellington record

Deleted: 7

A number of samples were identified in RLIMS (The GNS Science radiocarbon database) that were measured but not included in the dataset provided by Currie et al. (2011) and have been added to the record:

- NZA11524 was a “replacement” for NZA11278 which was discarded due to analytical problems (but these were collected on different dates).
- NZA11523 was a remeasurement of another extraction from bottle 12 (NZA9382) which was an obvious outlier in the record. Both samples had the same sample date, sampling bottle and $\delta^{13}C$.
- NZA8462-8468 are all extractions from the same sample bottle. They were originally given two different R numbers, and each was given a different NZA number. These have all been averaged together in this updated record.
- NZA26053 (bottle 96) and NZA39544 (bottle 42) are missing from the NIWA sample collection spreadsheet but are recorded in RLIMS. Collection dates can be guessed at, but not confirmed so have been flagged as having an analytical problem and been excluded from the dataset.
- NZA 11191, 11192, 11278, 39189 had analytical problems identified and have been flagged and removed.
- NZA11522 had a sampling problem and has also been flagged.
- Six gas counting samples were identified in RLIMS as contaminated and flagged (these were also excluded from the Currie et al. (2011) dataset).
- NZ 2284, 2290, 5700, 6535 were discovered in the RLIMS database that were not included in the Currie et al., (2011) dataset. These are not outliers and are added here.

9. Quality flags used in the datasets

Deleted: 8

A three-character flagging system is used. When no flag is indicated, each character is “.”, resulting in “…” for a result with no flags.

Deleted: three character

The first character is a "hard" flag. Any value in this column indicates a sampling or measurement problem and this result should be ignored in any analysis. These are retained in our internal database, but have been removed from the distribution dataset and may be obtained on request.

A.. Analytical measurement problem
N.. Problem at point of collection
C.. Contamination in sample
O.. This target was excluded as an outlier from the set of multiple targets measured for this sample

The second character is a "soft" flag, used to indicate that this result should be excluded from the background record analysis. Here it is used to indicate that the result is an outlier relative to the rest of the record, but there is no positively identified sampling or measurement problem and we therefore cannot remove the datapoint a priori. These values are included as a separate table in external distribution dataset but should not usually be included in any analysis.

.F. this measurement is a >3 sigma outlier from the smooth curve fit to the dataset
.A. period from 1990-1993 where the static NaOH samples deviate from recently measured flasks collected during the same time period and also deviate from the Cape Grim $^{14}CO_2$ measurements (Levin et al., 2010).
.T. period from 1995-2005 when the static NaOH samples have larger scatter than expected, and may be biased high, *but are retained in the record* in the absence of other data.

The third character is informational only and *does not indicate a problem with the result*. Here we use this flag to indicate changes from the results reported by Manning and Melhuish (1990) and/or Currie et al. (2011) and/or results initially reported from the Rafter AMS database to the NIWA database.

..D date of collection and/or decay correction has changed
..N normalization correction has changed
..A change in how the average of multiple targets from the same sample was calculated.
..O sample was excluded from multi-target average in Currie et al. (2011), but is now included in the analysis (applies only to the individual measurements, not listed for the mean values)

10. Data Availability

Deleted: 9

The datasets presented here are included in the supplementary material as a separate data file. Regularly updated versions of these datasets can be accessed through the World Data Centre for Greenhouse Gasses (http://ds.data.jma.go.jp/gmd/wdcgg/) or directly through GNS Science (https://gns.cri.nz/Home/Products/Databases/Wellington-atmospheric-14CO2-record) or NIWA (ftp://ftp.niwa.co.nz/tropac/).

11. Supplementary figures

Deleted: 9

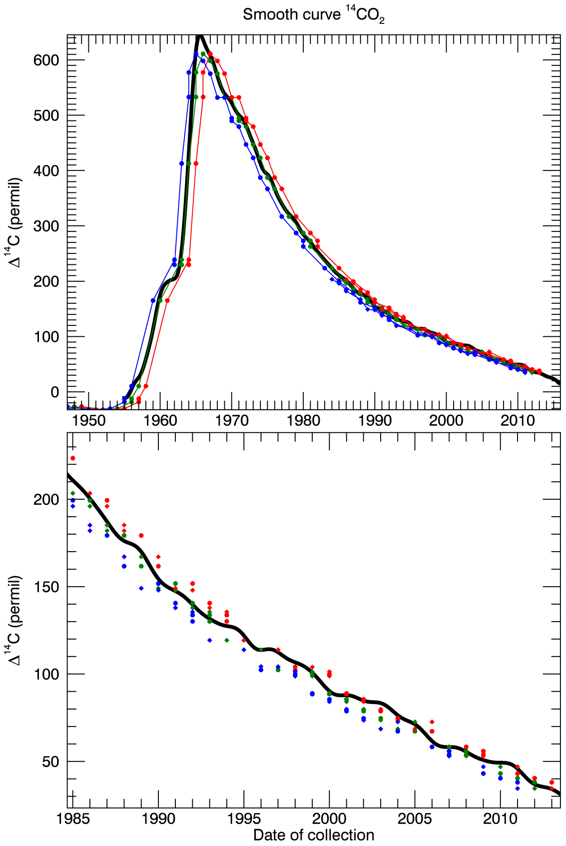

Figure S1. Evaluation of tree ring counts for the full record (top) and since 1985 (bottom). Black line is the smooth curve fit to the final Wellington $\Delta^{14}CO_2$ dataset. Green points and line are the combined tree ring data from both Baring Head and Eastbourne, using the assigned ring counts. Blue and red lines/points are the tree ring data shifted by one year in either direction. Adjustment in either direction clearly misaligns the upswing of the bomb pulse and the timing of the peak $\Delta^{14}CO_2$ values. In the later part of the record, since 1985, the trend through time is smaller, but the blue points (shifted one year older) consistently underestimate the Wellington $\Delta^{14}CO_2$ values throughout the record. The red points (shifted one year younger) overestimate until about 1995, when it is difficult to discern from this comparison whether the red or green (assigned ring count year) is correct, but there is excellent agreement between the records from four cores and three different trees, giving us confidence that the ring counts are correct in all cases.

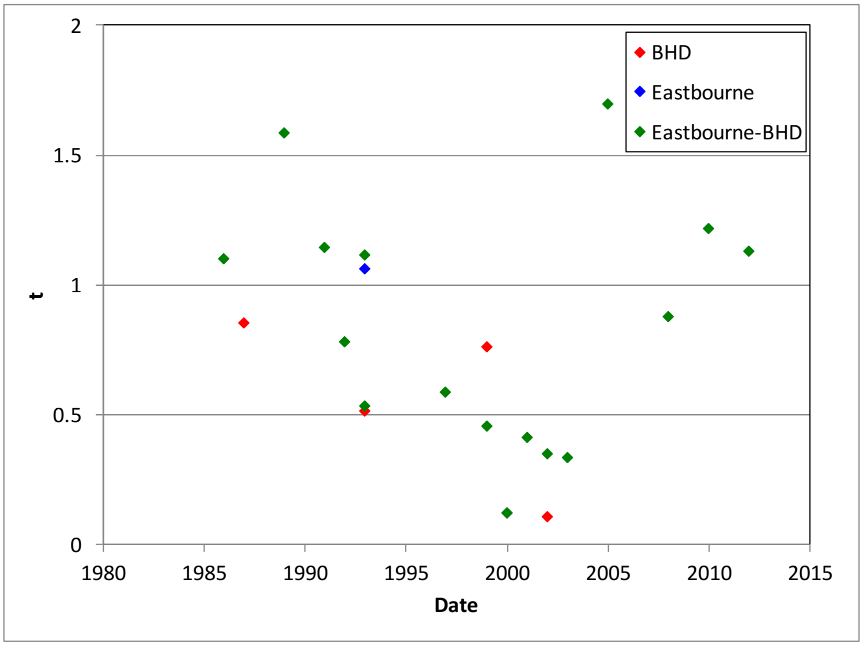

Figure S2. Evaluation of replicate tree ring measurements, shown as Student's t statistic for each replicate pair (pair difference divided by the sum of squares of the individual uncertainties). Red: replicates are the same ring taken from two different cores of the same pine tree at Baring Head. Blue: replicates are the same ring taken from cores of two different kauri trees from Eastbourne (located 20 m apart). Green: replicates are the same ring taken from the Baring Head pine and one Eastbourne tree. A t-value of 1 indicates that the two replicates agree within one standard deviation. Of the 22 replicate pairs, 14 (64%) agree within one standard deviation and 100% agree within two standard deviations. Using a paired (dependent) sample t-test to combine all the pairs of results, the overall t statistic is 0.97 (p=0.34), that is, given the assigned uncertainties, there is no significant difference between the pairs of measurements. There is no significant bias between the Baring Head and Eastbourne trees (mean difference is -0.4 ± 0.8 for the 16 replicate pairs that span both sites).