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 code is included in the dataset.

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

45 was variable). For samples measured by AMS (dates after 1995), there are only a few for
46 which the end date is missing:
47 • For NZA9240 (start date 1/4/96), NZA11523 (start date 15/6/96) and NZA11524
48   (start date 1/12/99). We assume a 14-day collection period since most other samples
49   during this period were collected over 14 days.
50 • NZA14572 (start date 11/9/00) and NZA14573 (start date 22/9/00). We assume the
51   end date is the same date as the start date for the following sample, which give days
52   exposed of 11 and 18 days respectively.
53

54 3. Collection methods

55 *3.1. Static NaOH absorption*

56 The primary collection method is static absorption of CO_2 into sodium hydroxide
57 (NaOH) solution, which is left exposed to air at the sampling site for ~2 weeks.
58 The NaOH solution is first prepared to remove any CO_2 that was initially present when
59 making up the solution. The standard preparation method involves bubbling CO_2-free air
60 or nitrogen through deionized water for ~5 hours. We continue using this step, although
61 it may be unnecessary since any initially-present CO_2 is removed in the following steps.
62 NaOH pellets are then added to make 0.5 or 1 M L^{-1}. The NaOH pellets always absorb
63 some ambient CO_2 when exposed to air, inevitably introducing some CO_2 contamination
64 during the preparation, and this is removed from the solution by the addition of ~ 1 g
65 barium chloride (0.004 mol) in ~50 mL of bubbled, deionized water. Any carbonate is
66 precipitated as barium carbonate and the resultant nominally carbonate-free solution is
67 decanted into smaller bottles under a stream of CO_2-free air or nitrogen. Since 2015, a
68 further step adds 0.2 mL of 85% phosphoric acid to the solution after the barium chloride
69 precipitation step. Any remaining barium is then precipitated as barium phosphate,
70 ensuring that all the CO_2 absorbed during sample collection remains as carbonate rather
71 than precipitating out as barium carbonate. The resulting solution is stored in sealed
72 bottles until ready for use.
73

74 From 1954-1995, ~ 2 L of the 0.5 M NaOH solution was poured into a large Pyrex® tray
75 and left exposed to air for 1 - 2 weeks. Some early (1954-1970) samples may have been
76 collected using different vessels. In a few early samples, air may have been pumped
77 through the solution rather than passive absorption (Manning et al., 1990). Two samples
78 were collected using Ba(OH)$_2$ rather than NaOH, in 1957 and 1958. As these are not
79 outliers, they are included in the dataset, with a comment and flag.
80

81 Beginning in 1995, when the measurement technique changed from gas counting to AMS
82 measurement (see section 3), the large trays were replaced with high density polyethylene
83 (HDPE) bottles containing ~200 mL 0.5 M NaOH solution, with the depth of liquid kept
84 the same as it was in the previously used trays. The lid is removed and the bottle is left
85 open inside a Stevenson meteorological screen for ~2 weeks. Date of opening and
86 closing bottle is recorded. The two methods (large tray and small bottle) have been
87 compared, with no significant difference observed between the two methods (Currie et al.,
88 2011).
89

90   Static NaOH absorption necessarily fractionates relative to CO_2 in the atmosphere;
91   typical $\delta^{13}C$ values are -15 to -25 ‰ for these samples. This is corrected for in the data
92   analysis.
93
94   *3.2. Whole air flask samples*
95   Flasks of whole air are collected by flushing ambient air through the flask for several
96   minutes. The flask is then closed and allowed to fill to slightly over ambient pressure.
97   All flask samples presented here were collected at the Baring Head site. Most of those
98   collected during 1984 – 1993 were collected during southerly, clean air, conditions
99   (Stephens et al., 2013). Flasks collected since 2013 were all collected during southerly
100 conditions.
101
102 *3.3. Tree rings*
103 When trees photosynthesize, they faithfully record the ^{14}C content of ambient CO_2 in
104 their cellulose, the structural component of wood. Annual tree rings therefore provide a
105 summertime (approximately September – April in the Southern Hemisphere) daytime
106 average $\Delta^{14}CO_2$. Photosynthetic uptake varies during the daylight hours depending on
107 factors including growth period, sunlight, and temperature (Bozhinova et al., 2013). For
108 the Wellington location with typically high wind speeds and little short-term variability in
109 CO_2 and other trace gases, this variability in photosynthesis is likely not a big impact on
110 $\Delta^{14}CO_2$. The ring age we assign differs from standard nomenclature for Southern
111 Hemisphere dendrochronology. We assign the mean age of the ring as January 1 of the
112 year in which growth finished (i.e. the mean age of a ring growing from September –
113 April), whereas dendrochronologists assign the "ring year" is as the year in which ring
114 growth started (i.e. the previous year).
115
116 We collected cores from three trees close to the Baring Head site. A pine (*Pinus radiata*)
117 located 10 m from the Baring Head sampling station (figure 1) yielded rings back to 1986
118 (Norris, 2015). New Zealand kauri (*Agathis australis*) is a long-lived hardwood species
119 that has been widely used in dendrochronology and radiocarbon calibration studies (Hogg
120 et al., 2013). The rings of this species are well defined and seldom missing, making this
121 an ideal species for our comparison. We collected cores from two specimens planted in
122 1919 and 1920, located 20 m from one another in Eastbourne, 12 km from Baring Head
123 (figure 1).
124
125 We evaluated the ring counts by the timing of the ^{14}C bomb spike. Shifting the
126 Eastbourne record by one year in either direction moves the maximum in the tree ring
127 record out of phase with the final Wellington $\Delta^{14}CO_2$ record (figure S1), confirming that
128 the ring counts are correct. For the Baring Head pine, rings go back to only 1986, so we
129 compare with the Eastbourne record, and find that the two records agree very well, with
130 an insignificant mean difference of -0.4 ± 0.8 ‰ (figure S1).
131
132 Sampling of the tree rings for ^{14}C analysis is also a potential source of bias; ideally only
133 the full ring for a given year would be sampled. In practice, it is difficult to ensure that
134 the full ring is isolated without losing any material from that ring and no wood from the
135 surrounding rings is included. To evaluate this, we measured replicate samples different

136   cores from the same tree (Baring Head pine) or two different trees (Eastbourne kauri).
137   For samples collected since 1985, all these replicates agree within one standard deviation
138   (figure S2). However, for three replicates from Eastbourne in 1963, 1965 and 1971, we
139   see large differences of 9.2, 44.5 and 4.9 ‰, which we attribute to small differences in
140   sampling of the rings which were magnified by the rapid change in $\Delta^{14}C$ of up to 200 ‰
141   yr^{-1} during this period. Thus, the ring values during this period should be treated with
142   caution.
143
144   ## 4. ^{14}C Analysis methods
145   ### 4.1. Extraction
146   From 1954- 1995 CO_2 was extracted from the NaOH solution by acidification followed
147   by cryogenic distillation (Rafter and Fergusson, 1959). AMS sample extraction follows
148   the general same method, but with smaller volumes (Currie et al., 2011), and this
149   extraction is performed at NIWA. The $\delta^{13}C$ is measured at NIWA on an aliquot of CO_2
150   from the same extraction.
151
152   CO_2 is extracted from whole air flask samples by cryogenic extraction. Samples
153   collected from 1984- 1993 were extracted at GNS Science by Graeme Lyon and archived
154   as ampoules of pure CO_2. In 2012, these tubes were cracked under vacuum to liberate the
155   CO_2. Any leakage during storage is readily by air present in the tube when it is cracked
156   for transfer and tubes with leakage were discarded. An aliquot of each was measured for
157    $\delta^{13}C$ to confirm that no fractionation had occurred during storage. Whole air samples
158   collected since 2013 are analyzed for $\delta^{13}C$ and other trace gases and isotopes at NIWA
159   and for the $^{14}CO_2$ measurement, CO_2 is extracted from whole air at GNS Science
160   (Turnbull et al. 2015).
161
162   Cellulose was isolated from whole tree rings by first removing labile organics with a
163   series of solvent washes (hexane, isopropanol and acetone) in a Soxhlet system (Norris,
164   2015). This was followed by oxidation to isolate the cellulose from other materials (Hua
165   et al., 2000). The cellulose was combusted and the CO_2 purified following standard
166   methods in the GNS Science laboratory (Baisden et al., 2013).
167
168   ### 4.2. Graphitization (AMS only)
169   The LG1 graphitization system was used from 1995 to 2011 (NZA < 50,000) (Lowe et al.,
170   1987), and replaced with the RG20 graphite system in 2011 (NZA > 50,000) (Turnbull et
171   al., 2015). In both systems, CO_2 is reduced to graphite over iron catalyst in the presence
172   of hydrogen gas. RG20 has online pressure monitoring to ensure the reaction goes to
173   completion. On LG1, iron powder was used as catalyst; on RG20 iron oxide is reduced
174   to iron online immediately prior to graphitization. All samples measured on the EN
175   Tandem were prepared using LG1, and all except one sample measured on XCAMS were
176   prepared using RG20.
177
178   ### 4.3. Decay counting
179   Static NaOH samples were measured by conventional decay counting on the CO_2 gas
180   from 1954-1995, described in detail in earlier publications (Manning and Melhuish, 1990;

181   Currie et al., 2011). Samples measured by gas counting are identified by unique numbers
182   with the prefix "NZ".

183

184   *4.4. AMS*
185   All measurements made since 1995, including recent measurements of flask samples
186   collected in the 1980s and 1990s, were measured by accelerator mass spectrometry
187   (AMS). The unique identifiers for these samples have the prefix "NZA", but the NZ and
188   NZA numbers never overlap.

189

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

201

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

207

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

213

214   First, including all measurements, we obtain χ^2_v =1.2 from 441 targets. We identified
215   outliers as those where χ^2_v for that sample date was greater than three and excluded the
216   targets within that set for which the residual was greater than two. Recalculating χ^2_v after
217   removing these outliers obtains χ^2_v=0.99 from 427 targets, a satisfactory result. From EN
218   Tandem measurements only, χ^2_v=1.02 from 397 targets. For XCAMS, multiple targets
219   were collected from both within and between wheels. The Baring Head NaOH samples
220   measured on XCAMS give, χ^2_v=0.95 from 102 targets. The weighted means have been
221   recalculated from the F^{14}C values, and Δ^{14}C calculated from F^{14}C.

222

223   For samples where the outliers changed, differences of up to 5‰ from the values reported
224   by Currie et al. (2011) occur. For other samples, differences of 0.1 – 0.5 ‰ occur,
225   apparently because when results were initially calculated, straight rather than weighted

226 means (weighted by reported uncertainties) were used in some cases and rounding errors
227 in other cases.
228
229 A few samples were measured on the EN Tandem in two different wheels, and in some
230 cases were assigned different R (laboratory) numbers. In those cases, Currie et al. (2011)
231 did not always average all four replicates. This has now been revised.
232
233 In 2010, the EN Tandem was replaced with a National Electrostatics Corporation AMS,
234 dubbed XCAMS (NZA > 34,000). XCAMS measures all three carbon isotopes, such that
235 the normalization correction is performed using the AMS measured ^{13}C values
236 (Zondervan et al., 2015). XCAMS measurements are made on single graphite targets,
237 measured to high precision, typically better than 2‰ overall uncertainty.   Replicate
238 analyses within and between wheels have been used to determine long-term repeatability
239 and overall uncertainty levels (Turnbull et al., 2015), including replicate measurements of
240 samples from this dataset. Where more than one measurement was made for a given date,
241 we report the weighted mean of all measurements.
242
243 Three different preparation methods have been used for the primary standards (oxalic
244 acid I, OxI) using in determining $\Delta^{14}CO_2$. Sealed tube combustion was used for all the
245 EN Tandem measurements, whereby sufficient solid OxI material was combusted to
246 provide five or six aliquots. The resultant CO_2 was split and then graphitization was
247 performed separately for each aliquot. Since 2010, OxI has also been prepared by
248 combustion of single aliquots by elemental analyzer and subsequent cryogenic CO_2
249 collection (Baisden et al., 2013). Both these methods show larger scatter than air
250 standard materials, and there is an offset of -1.5‰ in samples measured with OxI made
251 by sealed tube combustion, and those samples have been corrected for this offset
252 (Turnbull et al., 2015). To reduce the scatter in OxI due to variability in combustion
253 (which air samples do not undergo), we prepared two large flasks of CO_2 from OxI, and
254 in 2015, began taking aliquots of this CO_2 as primary standard material. There is no
255 offset relative to the elemental analyzer combustion method, and repeatability is slightly
256 improved.
257
258 During the transition between the EN Tandem and XCAMS instruments, three
259 Wellington $^{14}CO_2$ samples were measured on both instruments, and XCAMS results were
260 4.2‰ higher on average. However, this jump during the transition does not appear
261 consistent across wheels, and is not apparent in the larger quality control dataset (Baisden
262 et al., 2013), and is not addressed further here.
263

264 ## 5. Results format

265 ### 5.1. Date of collection
266 NaOH samples are collected over a period of typically two weeks, and sometimes much
267 longer. We determine the date of collection as the average of the start and end collection
268 dates. In cases where the end date was not recorded, we use the start date. For a few
269 samples, the sampling dates were not recorded or are ambiguous, and those results have
270 been flagged and excluded from the final dataset.
271

272   *5.2 $\Delta^{14}C$ and $F^{14}C$*

273   Results are reported here in $F^{14}C$ (Reimer et al., 2004) and $\Delta^{14}C$ (Stuiver and Polach,

274   1977). $F^{14}C$ is corrected for isotopic fractionation and blank corrected. $F^{14}C$ is

275   calculated from the original measurement data recorded in our databases, and has been

276   updated for only a handful of records where errors were found in the $\delta^{13}C$ value

277   originally used in the normalization correction. No changes in measured values were

278   made except in cases where multiple targets were averaged.

279

280     $\Delta^{14}C$ is derived from $F^{14}C$, with a correction for radioactive decay since the time of

281   collection (Stuiver and Polach, 1977). The date of collection is required for this

282   calculation. In previously published versions of the dataset, the gas counting and EN

283   Tandem results used only the year of collection in the $\Delta^{14}C$ calculation, and in many

284   cases, $\Delta^{14}C$ was initially reported using the date of measurement. Currie et al. (2011) re-

285   calculated $\Delta^{14}C$ in a separate database, using the date of collection, but assumed that the

286   date of measurement, rather than the year of measurement, had been used. Here we went

287   back to the original measured $F^{14}C$, and to the sample collection database to find the date

288   of collection. All $\Delta^{14}C$ results are recalculated using the decimal date of collection. This

289   results in changes of a few tenths of permil in most $\Delta^{14}C$ values relative to those reported

290   by Currie et al (2011) and Manning and Melhuish (1990). Samples for which changes

291   have been made relative to the previously published results are indicated by the third

292   column quality flag provided in the supplementary dataset.

293

294   *5.3. Uncertainties*

295   The primary source of uncertainty is the ^{14}C counting statistical uncertainty determined

296   from the number of beta decays (gas counting) or ^{14}C counts (AMS). For AMS

297   measurements, we add an additional error term, determined from the long-term

298   repeatability of secondary standard materials and added in quadrature to the AMS

299   uncertainty. In early AMS measurements, this was 4%, and has decreased through time to

300   0.12%. For the EN Tandem results, this was determined from the performance of OxI

301   primary standard targets through time, which may somewhat overestimate the uncertainty

302   for air samples (Turnbull et al., 2015). For XCAMS measurements, long-term

303   repeatability of air samples has been assessed based on repeated splits of CO_2 from a

304   subset of these NaOH samples, as well as repeated extractions of aliquots of CO_2 from

305   pressurized tanks of whole air. These repeated aliquots are measured both within the

306   same AMS wheel and across multiple wheels. Within individual wheels, the counting

307   statistical uncertainty is sufficient to explain the full variability, but an additional

308   uncertainty of 0.12% is required to explain the spread across multiple wheels (Turnbull et

309   al., 2015).

310

311   ## 6. Data validation

312   *6.1. Tree ring comparison*

313   Over the more than 60 years of the Wellington $\Delta^{14}CO_2$ record, there have necessarily

314   been many changes in methodology, and the tree rings provide a way to validate the full

315   record, albeit with lower resolution. Due to the possible sampling biases in the tree rings,

316   use them to validate the existing measurements, but do not include them in the main

317   Wellington $\Delta^{14}CO_2$ dataset. We include the tree ring data in a separate table for
318   researchers who may want to use these data in addition to the main Wellington $\Delta^{14}CO_2$
319   dataset.
320
321   During the rapid $\Delta^{14}CO_2$ change in the early 1960s, there are some differences between
322   the kauri tree ring and Wellington $\Delta^{14}CO_2$ records. The 1963 and 1964 tree ring samples
323   are slightly lower than the concurrent $\Delta^{14}CO_2$ samples. The peak $\Delta^{14}CO_2$ measurement
324   in the tree rings is 30‰ lower than the smoothed $\Delta^{14}CO_2$ record, and 100‰ lower than
325   the two highest $\Delta^{14}CO_2$ measurements in 1965. These differences are likely due to small
326   errors in sampling of the rings, which will be most apparent during periods of rapid
327   change.
328
329   Prior to 1960 and from the peak of the bomb spike in 1965 until 1990, there is remarkable
330   agreement between the tree rings and Wellington $\Delta^{14}CO_2$ record, with the wiggles in the
331   record replicated in both records. And since 2005, there is excellent agreement across all
332   the different records. Some differences are observed in 1990-1993 and 1995-2005, which
333   we discuss in the following sections.
334

335   *6.2. 1990-1993 excursion*
336   An excursion in the gas counting measurements between 1990 and 1993 has previously
337   been noted as a deviation from the Cape Grim $\Delta^{14}CO_2$ record (Levin et al., 2010) during
338   the same period. Cape Grim is at a similar latitude and also draws air from Southern
339   Ocean. The Wellington and Cape Grim records overlap during almost all other periods.
340
341   We use archived CO_2 from flask samples to evaluate this period of deviation. First, the
342   recent flask samples agree very well with the NaOH static samples from the same period
343   (figure 2). This indicates that despite the difference in sampling period for the two
344   methods, flask samples reflect the $\Delta^{14}CO_2$ observed in the longer-term NaOH static
345   samples. We selected a subset of the archived 1984 - 1992 CO_2 samples for
346   measurement, mostly from Southerly wind conditions, but including a few from other
347   wind conditions. Each sample was liberated and the $\delta^{13}C$ measured before $\Delta^{14}CO_2$
348   measurement was performed. A few of the samples had been compromised, and this was
349   readily apparent as they contained air rather than CO_2. $\delta^{13}C$ from all the remaining
350   samples was in agreement with $d^{13}C$ measured from separate flasks collected at Baring
351   Head and measured for $\delta^{13}C$ by Scripps at close to the time of collection. These flask
352    $\Delta^{14}CO_2$ measurements do not exhibit the excursion seen in the NaOH static samples,
353   implying that the deviation observed in the original NaOH static samples may be a
354   sampling error. Annual tree rings from both the kauri and pine follow the flask
355   measurements for this period (figure 2), confirming that the NaOH static samples are
356   anomalous.
357
358   The 1990-1993 period was characterized by major changes in New Zealand science, both
359   in the organizational structure and personnel. Although we are unable to exactly
360   reconstruct events that time, we hypothesize that the NaOH solution preparation was
361   conducted slightly differently, perhaps omitting the $BaCl_2$ precipitation step for these
362   samples. This would result in contaminating CO_2 absorbed on the NaOH before the

363 solution was prepared, which would result in higher $\Delta^{14}CO_2$ observed in these samples
364 than in the ambient air. In any case, these values are anomalous and we remove the
365 original NaOH static sample measurements between 1990 and 1993 and replace them
366 with the new flask measurements for the same period.
367

6.3. 1995-2005 variability

369 In 1995, the measurement method was changed from gas counting to AMS. During the
370 first ten years of AMS measurements, the record is much noisier than during any other
371 period (figure 2). In 2005, online ^{12}C measurement was added to the AMS system,
372 substantially improving the measurement accuracy, and the noise in the $\Delta^{14}CO_2$ record
373 immediately reduced.
374

375 The remaining NaOH solution for all samples collected since 1995 has been archived,
376 and typically only every second sample collected was measured, with the remainder
377 archived without sampling. In 2011-2016, we revisited the 1995-2005 period,
378 remeasuring some samples that had previously been measured and some that had never
379 been measured for a total of 52 new analyses.
380

381 The new measurements on this period do show reduced scatter over the original analyses,
382 particularly for the period from 1998-2001 where the original analyses appear
383 anomalously low and in 2002-2003 when the original analyses appear anomalously high.
384 Yet there remain a number of both low and high outliers in the new measurements.
385 These are present in both the samples that were remeasured and in those for which this
386 was the first sample from the bottle. This suggests that a subset of the archived sample
387 bottles were either contaminated at the time of collection, or that some bottles were
388 insufficiently sealed, causing contamination with more recent CO_2 during storage.
389 Comparison with the tree ring measurements and with the Cape Grim record (Levin et al.,
390 2010) suggest that the measurements during this period may, on average, be biased high
391 as well as having additional scatter (figure 3). Nonetheless, in the absence of better data,
392 we retain both the original and remeasured NaOH sample results in the full record.
393

7. Existing results added to the Wellington record

395 A number of samples were identified in RLIMS that were measured but not included in
396 the dataset provided by Currie et al. (2011) and have been added to the record:
397 • NZA11524 was a "replacement" for NZA11278 which was discarded due to
398   analytical problems (but these were collected on different dates).
399 • NZA11523 was a remeasurement of another extraction from bottle 12 (NZA9382)
400   which was an obvious outlier in the record. Both samples had the same sample date,
401   sampling bottle and $\delta^{13}C$.
402 • NZA8462-8468 are all extractions from the same sample bottle. They were originally
403   given two different R numbers, and each was given a different NZA number. These
404   have all been averaged together in this updated record.
405 • NZA26053 (bottle 96) and NZA39544 (bottle 42) are missing from the NIWA sample
406   collection spreadsheet but are recorded in RLIMS. Collection dates can be guessed at,
407   but not confirmed so have been flagged as having an analytical problem and been
408   excluded from the dataset.

409   •   NZA 11191, 11192, 11278, 39189 had analytical problems identified and have been
410   flagged and removed.
411   •   NZA11522 had a sampling problem and has also been flagged.
412   •   Six gas counting samples were identified in RLIMS as contaminated and flagged
413   (these were also excluded from the Currie et al. (2011) dataset).
414   •   NZ 2284, 2290, 5700, 6535 were discovered in the RLIMS database that were not
415   included in the Currie et al., (2011) dataset. These are not outliers and are added here.
416

417 ## 8. Quality flags used in the datasets

418 A three character flagging system is used. When no flag is indicated, each character is
419 ".", resulting in "…" for a result with no flags.
420

421 The first character is a "hard" flag. Any value in this column indicates a sampling or
422 measurement problem and this result should be ignored in any analysis. These are
423 retained in our internal database, but have been removed from the distribution dataset and
424 may be obtained on request.
425 A..   Analytical measurement problem
426 N..   Problem at point of collection
427 C..   Contamination in sample
428 O..   This target was excluded as an outlier from the set of multiple targets measured
429   for this sample
430

431 The second character is a "soft" flag, used to indicate that this result should be excluded
432 from the background record analysis. Here it is used to indicate that the result is an
433 outlier relative to the rest of the record, but there is no positively identified sampling or
434 measurement problem and we therefore cannot remove the datapoint a priori. These
435 values are included as a separate table in external distribution dataset but should not
436 usually be included in any analysis.
437 .F.   this measurement is a >3 sigma outlier from the smooth curve fit to the dataset
438 .A.   period from 1990-1993 where the static NaOH samples deviate from recently
439 measured flasks collected during the same time period and also deviate from the Cape
440 Grim $^{14}CO_2$ measurements (Levin et al., 2010).
441 .T.   period from 1995-2005 when the static NaOH samples have larger scatter than
442 expected, and may be biased high, *but are retained in the record* in the absence of other
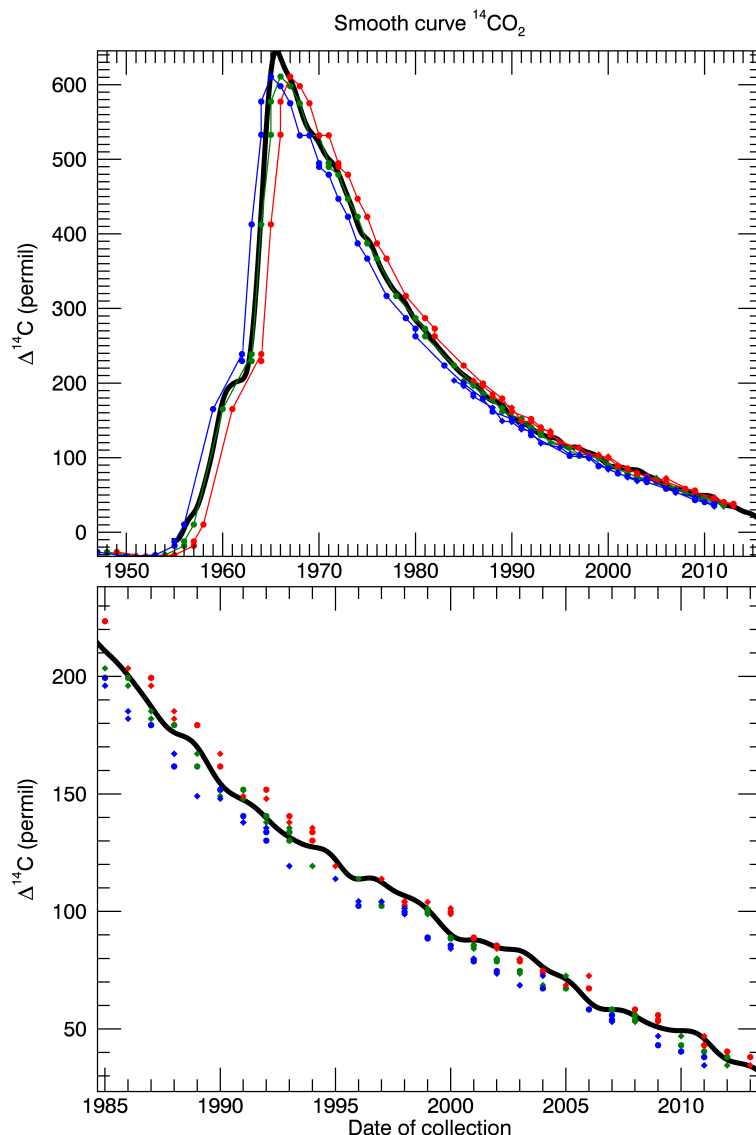443 data.
444

445 The third character is informational only and *does not indicate a problem with the result*.
446 Here we use this flag to indicate changes from the results reported by Manning and
447 Melhuish (1990) and/or Currie et al. (2011) and/or results initially reported from the
448 Rafter AMS database to the NIWA database.
449 ..D   date of collection and/or decay correction has changed
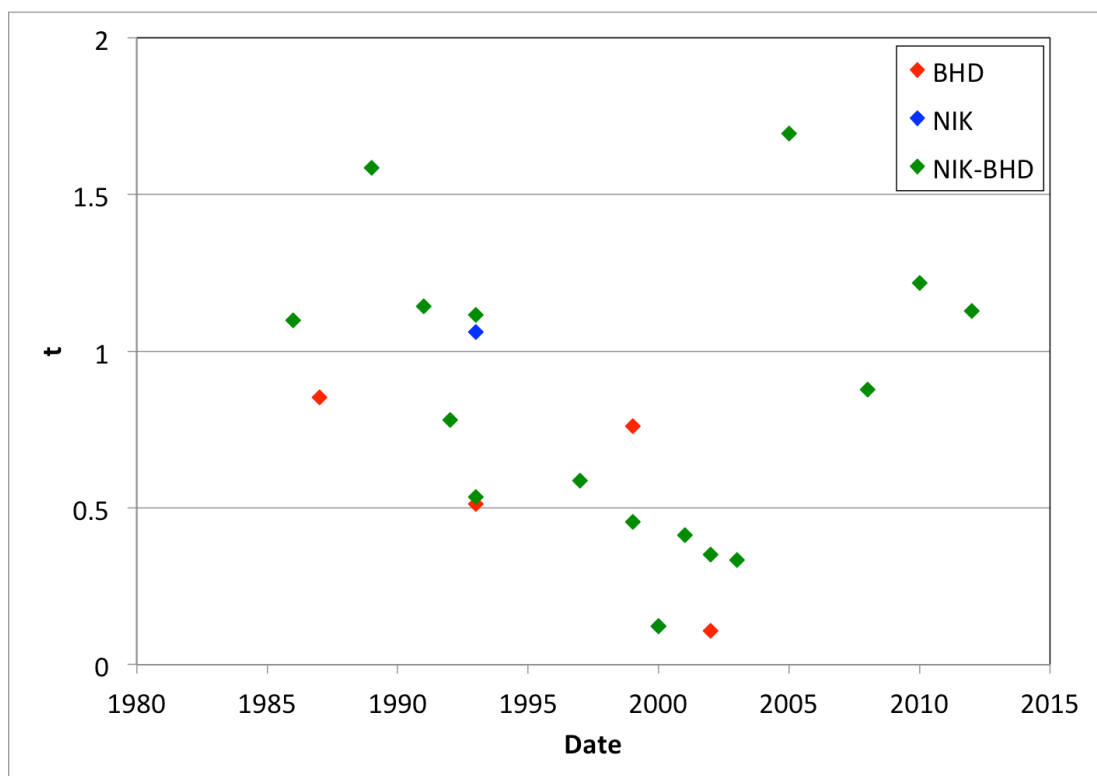450 ..N   normalization correction has changed
451 ..A   change in how the average of multiple targets from the same sample was
452 calculated.

453 ..O sample was excluded from multi-target average in Currie et al. (2011), but is now
454 included in the analysis (applies only to the individual measurements, not listed for the
455 mean values)
456

Figure S1. Evaluation of tree ring counts. Black line is the smooth curve fit to the final Wellington $\Delta^{14}CO_2$ dataset. Green points and line are the 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.

471

472
473 Figure S2. Evaluation of replicate tree ring measurements, shown as Student's t statistic
474 for each replicate pair. Red: replicates are the same ring taken from two different cores
475 of the same pine tree at Baring Head. Blue: replicates are the same ring taken from cores
476 of two different kauri trees from Eastbourne (located 20 m apart). Green: replicates are
477 the same ring taken from the Baring Head pine and one Eastbourne tree. A t-value of 1
478 indicates that the two replicates agree within one standard deviation. Of the 22 replicate
479 pairs, 14 (64%) agree within one standard deviation and 100% agree within two standard
480 deviations. There is no significant bias between the Baring Head and Eastbourne trees
481 (mean difference is -0.4 ± 0.8 for the 16 replicate pairs that span both sites).