Southern Hemisphere tree-rings as proxies to reconstruct Southern Ocean upwelling

Christian Lewis[1], Rachel Corran[1], Sara E. Mikaloff-Fletcher[3], Erik Behrens[3], Rowena Moss[3], Gordon Brailsford[3], Andrew Lorrey[3], Margaret Norris[1], Jocelyn Turnbull[1,2]

[1]GNS Science, Rafter Radiocarbon Laboratory, Lower Hutt, New Zealand
[2]CIRES, University of Colorado at Boulder, Boulder, Colorado, USA
[3]National Institute of Water and Atmospheric Research (NIWA), Wellington, New Zealand

Correspondence to: Dr. Christian Lewis (c.lewis@gns.cri.nz)

Contained in this Supplementary Information is the following:

- Further explanation of the Monte Carlo scheme to estimate uncertainties in data interpolation
- Discussion of GLODAP and HYSPLIT data analysis
- Exploration of a difference between two sites on Isla Navarino, Chile
- Documentation of the scripts used to produce the output for this manuscript
- Description the Open Access Data submitted alongside the manuscript
- Tree-ring validation figures
- HYSPLIT back-trajectory heatmaps for all sites.

S1. Monte Carlo Uncertainty Estimation of Interpolated Data

All tree-ring $\Delta^{14}C$ measurements are associated with dates centered in the summer of the growth season (i.e., for a tree growing between Southern Hemisphere spring to Autumn, September to May for example, the midpoint is peak summer; January 1). We want these measurements to be contextualized by the Southern Hemisphere Background (SHB) reference, a combination of two long-term records originating from University of Heidelberg's Cape Grim station, and GNS/NIWA Baring Head station. However, in order to find the difference between tree-ring $\Delta^{14}C$ and the SHB, the temporal axes need to be matched.

To achieve this, we smooth the SHB using use the NOAA Global Monitoring Laboratory's CCGCRV curve fitting method (Thoning et al., 1989), setting the algorithm to output data to match a concatenated list of the SHB collection dates and tree-ring dates. Both the "smooth" and "trend" functions are employed, however, only the "trend" function output is used in the final manuscript. To estimate the uncertainty the output, smoothed SHB, the CCGCRV algorithm is run inside a Monte Carlo loop. It proceeds as follows (line numbers refer to X_my_function.py; https://github.com/christianlewis091/science_projects/blob/main/SOAR_Tree_rings/scripts_OPEN_ACCESS/X_my_functions.py:)

1. Initial data is fed into the loop, including a) SHB x-values, b) x-values where output will be assigned (the concatenated SHB and tree-ring x-values), c) SHB y-values ($\Delta^{14}C$, ‰) d) SHB y-value error ($\Delta^{14}C$, ‰), e) parameterization of the FFT cutoff, d) times to loop (10k).
2. The first for-loop (lines 311-326):
    a. Iterate through the y-values (SHB $\Delta^{14}C$'s), and randomly return a value within the normal distribution of that y-value's error-range.
    b. Create "n" (10,000) sets of the randomized SHB, all stacked up
3. Second for-loop (lines 338-341):
    a. Iterate through the stack (array) from the previous loop (each iteration is a randomized SHB, referred to below as "sub-SHB")
    b. Run the sub-SHB through ccgFilter (line 340).
    c. Save the smoothed-output in a vertical stack (line 341).
4. Third for-loop (lines 354-364):
    a. For each x-value, find the mean and standard deviation of the smoothed outputs (a mean and standard deviation for smoothed y-values for each individual x-value in time). This is the value used for the remainder of analysis.

S2. GLODAP and HYSPLIT Data Analysis

Analysis of the GLODAP and HYSPLIT both required the interpolation of canonical Southern Ocean front data (Orsi, 1995) to longitude values consistent with a) the GLODAP DIC $\Delta^{14}C$ measurements and 2) HYSPLIT back trajectory data points. In other words, to know if a GLODAP $\Delta^{14}C$ data-point, or a HYSPLIT back-trajectory temporal snapshot is in the Antarctic Southern Zone, I must know the latitudes of those fronts at the exact longitude to compute it's "region".

The interpolation was performed using the "numpy.interp" function, and was verified to be well-constructed by visually looking at the interpolation over the original (Orsi, 1995) points. This can be seen in **Supplementary Figure 1**, with dotted lines showing (Orsi, 1995) fronts, and red scattered points showing the interpolated front at a new given longitude.

After interpolations shown in Supplementary Figure 1 were complete, data was parsed into different Southern Ocean zones by comparing the actual latitude to frontal latitudes. This was verified to be working by visual inspection. Examples of parsed GLODAP and HYSPLIT outputs are shown in **Supplementary Figure 2.** After verification of successful binning, the data can be averaged more easily using python's pandas library.

These codes can be found below. For HYSPLIT code, see HYSPLIT_check_Dec3_2024.py. For GLODAP code see, GLODA_check_Nov29_2024. The final data can be found in the "HYSPLIT" and "GLODAP" tabs of the associated data file.
https://github.com/christianlewis091/science_projects/tree/main/soar_tree_rings/scripts_EGU_submission

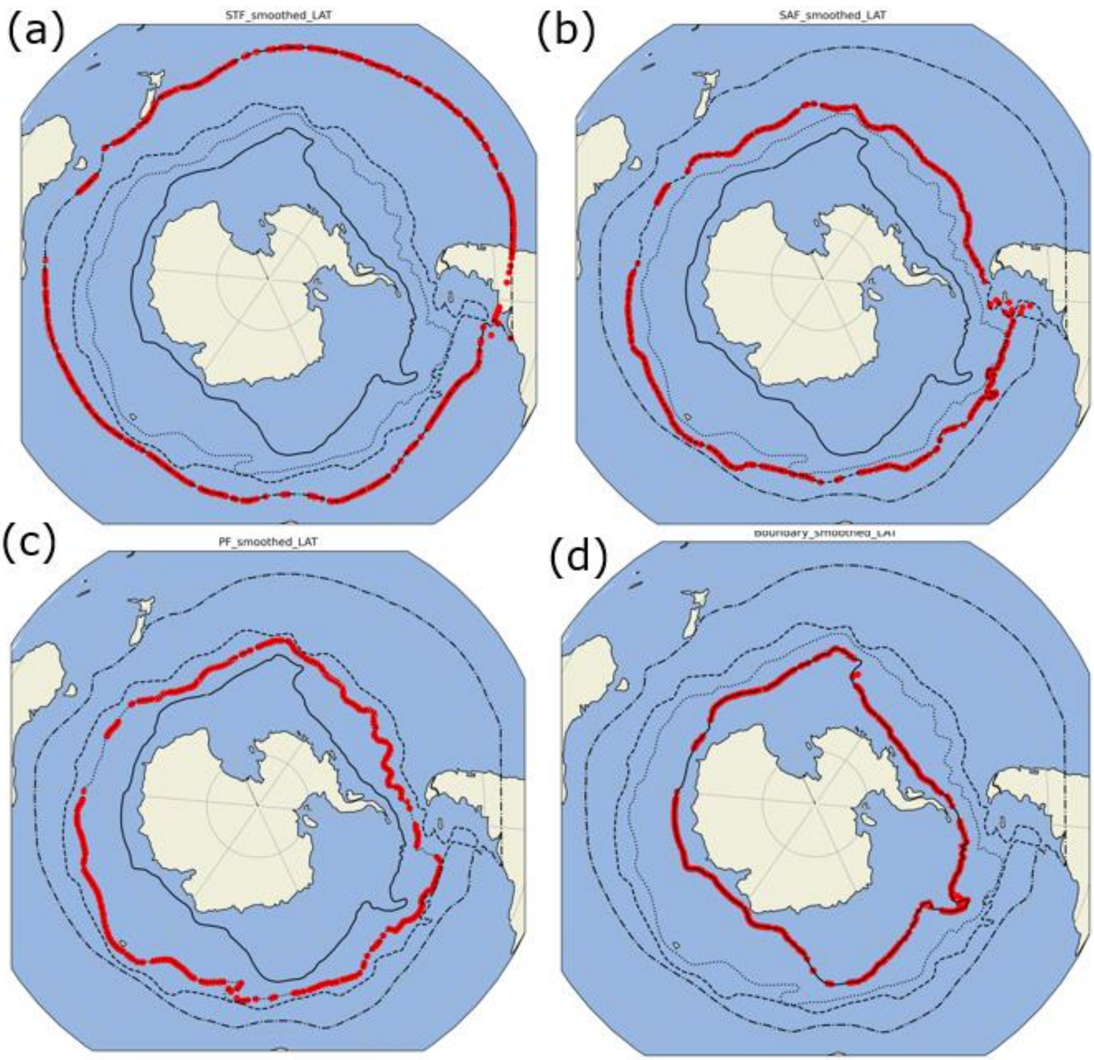

Fig. S1. (Orsi, 1995) fronts (dotted lines) overlaid with interpolated front at a new given longitude (red scattered points). The scattered data represent longitudes of GLODAP surface ocean DIC $\Delta^{14}C$ measurements

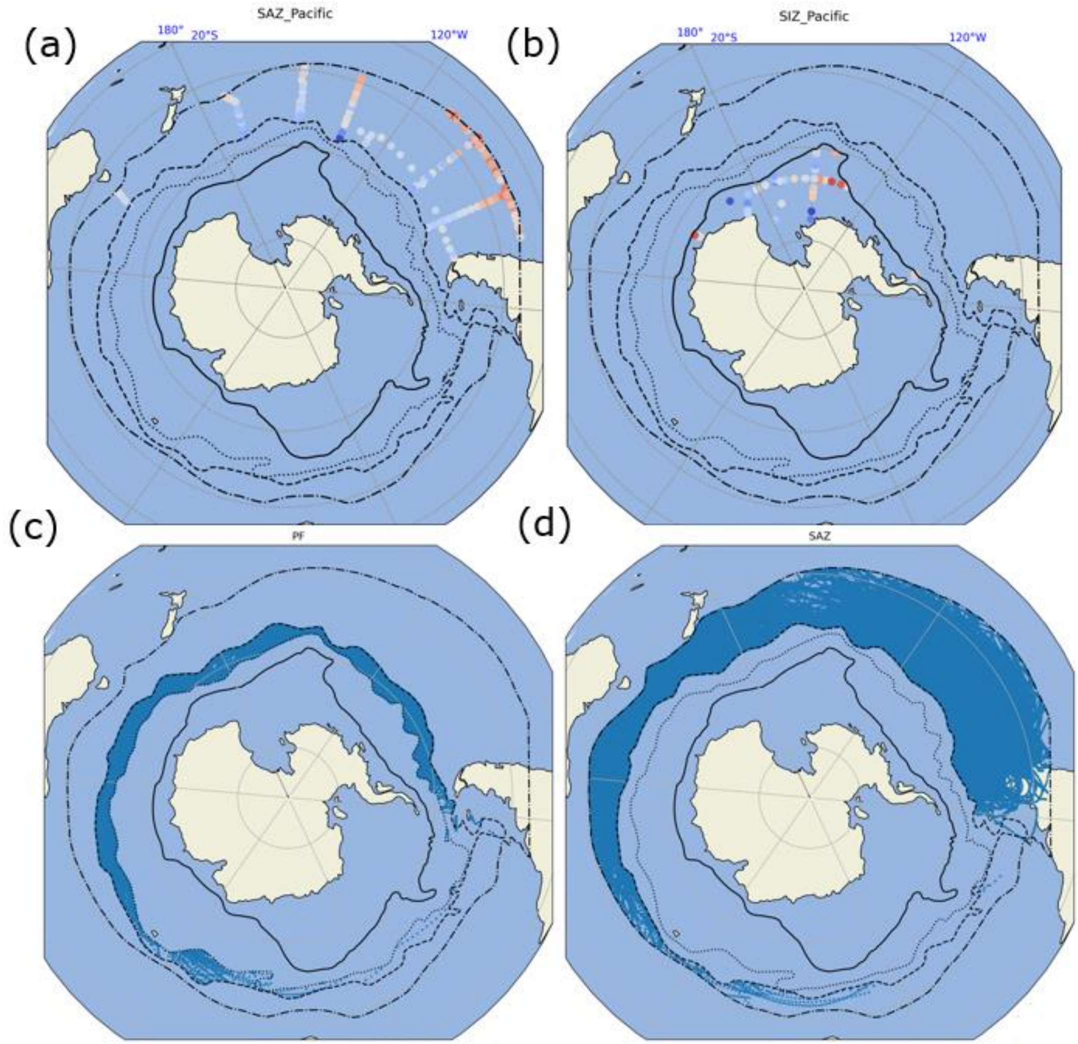

Fig. S2. Examples of visual verification that data has been properly binned between the Southern Ocean frontal zones. A) and B) show GLODAP measurements binned to the Subantarctic Zone (SAZ) and the Seasonal Ice Zone (SIZ), respectively. C) and D) show HYSPLIT output binned into the Polar Frontal Zone (PFZ) and Subantarctic Zone (SAZ). Remaining examples added to Supplementary_Figures.pptx

S3. Is there an observable difference between two sites on Isla Navarino, Chile?

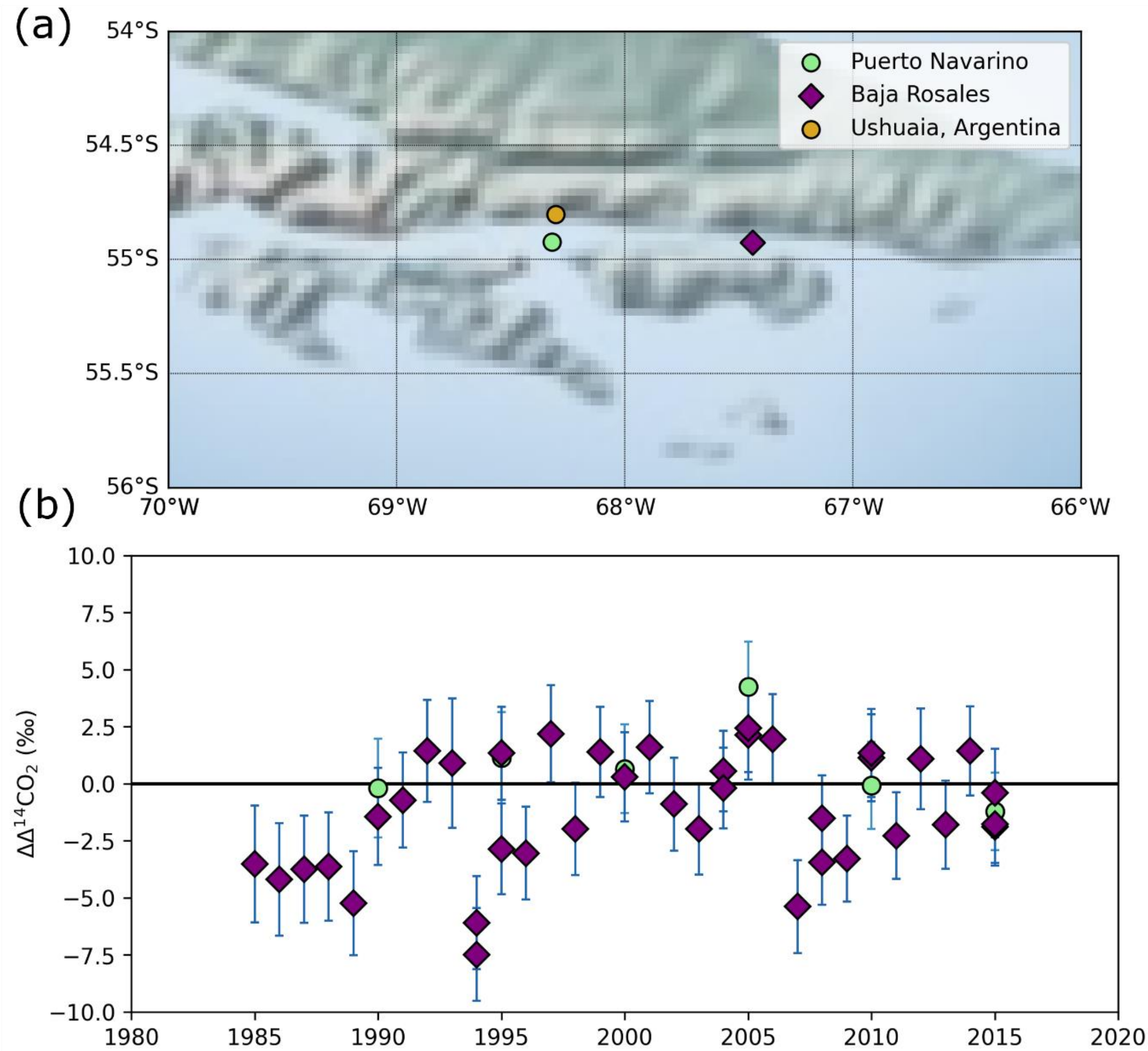

Fig. S3. Mean $\Delta\Delta^{14}CO_2$ of Puerto Navarino and Baja Rosales are 0.7± 1.7‰ and -1.2±2.5‰, respectively. An independent t-test yields a p-value of 0.08, which is too high to reject the null-hypothesis that the data are not different.

S4. Description and location of python files used to produce output for this work (for sharing but also personal future reference)

All the files described below can be found in the following GitHub directory. https://github.com/christianlewis091/science_projects/tree/main/soar_tree_rings/scripts_EGU_submission

Please refer to the following bullets for a brief description of what each script is doing. These directories were cleaned and made sure that they run properly before submission on Dec 4, 2024. Previous versions of the code are version controlled, and for further questions, please reach out to the corresponding author.

These codes were run in Python 3.9, and dependencies are listed in Dependencies.txt in the above Github folder.

S4.1. ***Preparing Reference Background***

1. **Reference2:** In a previous unpublished work, a slight offset between the Rafter Radiocarbon Lab and Heidelberg University was found between 1986-1994. This script applies an offset correction for those 8 years and merges the two time-series together. This record becomes one of the two considered for use as a background with which to compare the tree ring measurements, but in the end is not used. In the final version of the work, and OPEN_ACCESS_DATA3.xlsx , this is referred to as **Reference 2 or Ref2.**
2. **Reference1:** This file is similar to **B_CGO_BHD_harmonization,** in that is merges the two datasets, but in this case, no offset correction is applied. In the final version of the work, and OPEN_ACCESS_DATA3.xlsx , this is referred to as **Reference 1 or Ref1.**

S4.2. ***Cleaning Raw Data***

3. **Tree_ring_analysis.py:** This script plots our tree-ring $\Delta^{14}C$ measurements on top of the harmonized background record created in **Reference1.** We look for obvious false ring counts, as described in the methods section, and remove bad cores from the record. This sheet outputs a “flagged” and “clean” datasheet for reference. These are contained in the OPEN_ACCESS_DATA3.xlsx file.
4. **August 1, 2024:** The document that is read into this script (below) was manually edited to include three wheels with tree ring data run by Pene. These were in TW3516, 3519, and 3522. From these data (and analyses in **Tree_Ring_Second_Check.py:**, we concluded that Raul Main should be excluded completely because it was a miscount (RC’s old count was clearly offset from Pene’s, which fall in line with the rest of the data.) This new compiled data was re-run through the scripts flown below.
   df = pd.read_excel(r'H:\Science\Datasets\SOARTreeRingData2022-02-01_August_1_2024.xlsx')

S4.3. ***Data Manipulation***

1. **Reference_to _sample _xvals.py:** I need to see differences from the backgrounds at each X-value for samples. If the sample x-values don’t match the reference x-values, I can’t subtract them. This sheet aims to take the data from ”sampleset merge.py” and create CCGCRV smoothed and trended output at those x-values for the two types of

reference datasets we have. After the reference outputs are created, they need to be re-merged with the sample data. This actually caused a lot of problems in pandas, since many of the tree rings have multiple measurements for the same date, it caused a lot of scrambling. I figured out a solution, but it includes a slightly longer for loop. To keep things simple. I limit this script to only the production of the output of the Monte Carlo simulation that gets means of output at sample x-values.

2. **Reference_to _sample_xvals2.py:** This script takes the output from the script above, and uses a for loop to match dates for the samples, to dates in the monte carlo output. Based on matching dates, the reference data is appended to the sample dataframe and now the two can be compared more simply.

S4.4. *Data Analysis*

1. **Main_analysis.py:** After May 1, 2023 when we decided to simplify the first paper and push all ocean model comparisons to the future, I could simplify this paper to one file. That's all in here.
2. **HYSPLIT_check_Dec3_2024:** makes Hysplit heatmap heatmaps, calculates the time each site spends in each ACC zone. Further described in the section above
3. **GLODAP_tidy.*py*:** Reads and tidys up some GLODAP data from complete merged file.
4. **GLODAP_check_Nov29_2024:** makes the figure that shows nitrate and D14C on the orthographic Earth.
5. **Map_function.py:** A small file of mapping functions for GLODAP script.

S5. Complete Record of Ring-Count Validation

Ensuring that tree-rings are counted properly for correct chronologies is critical. This is described in the final paragraph of the main manuscript section 2.1. Below are figures used for ring-count validation for each site. Additional per-site information and descriptions are in figure captions.

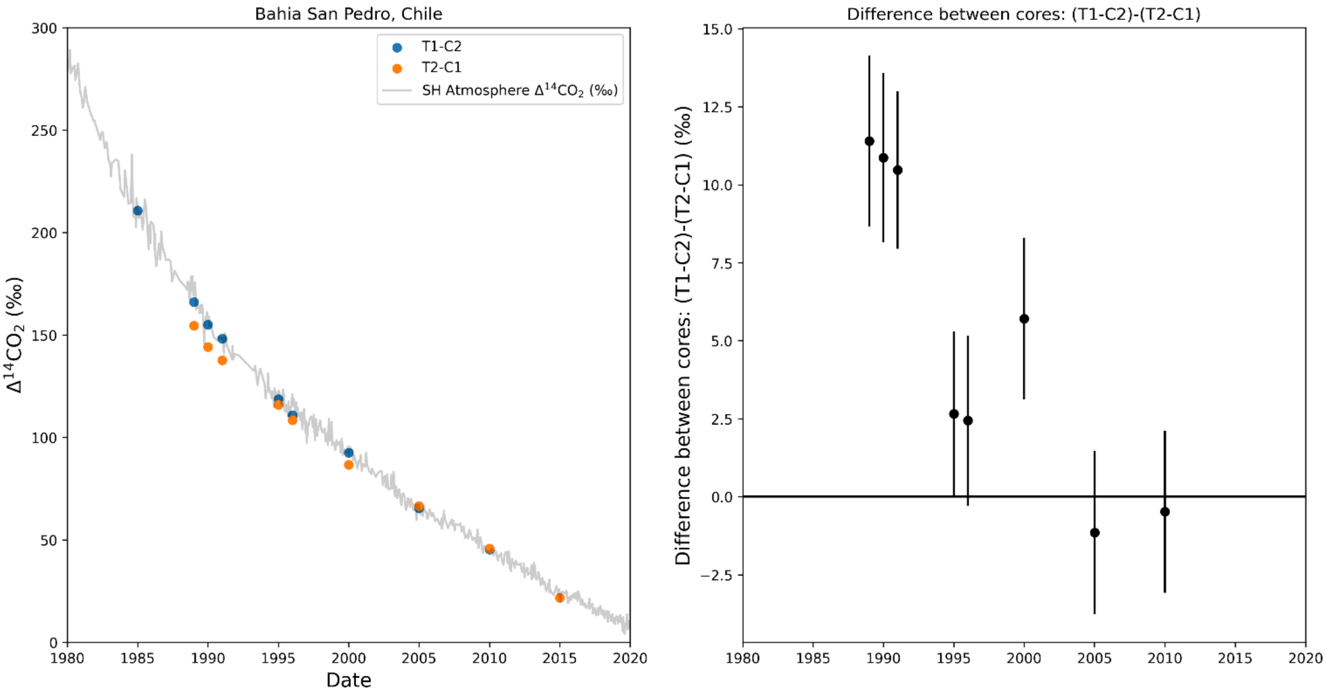

Fig. S4. The Bahia San Pedro record includes two trees, each with multiple cores. The second core from Tree 1 and first core from Tree 2 were chosen for measurement. The two records deviate from each other before 2005, therefore, all data before 2005 has been removed from the analysis.

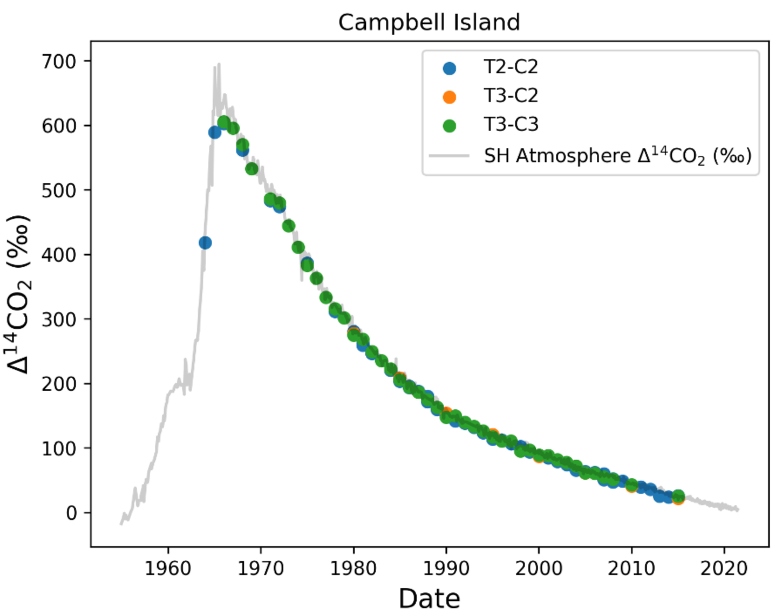

Fig. S5. No data was removed from Campbell Island Record.

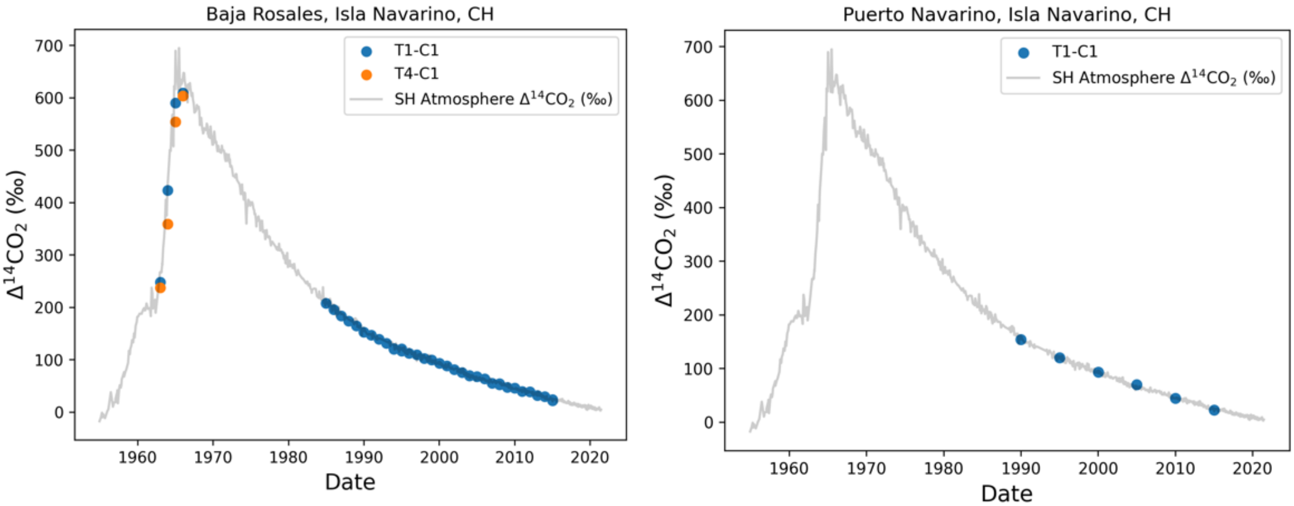

Fig. S6. Of the two neighboring sites, Puerto Navarino lies further west and is in proximity to the Argentinian city of Ushuaia, while Baja Rosales is to the east. These two sites were selected with the expectation that any significant land biosphere signal or fossil fuel emissions from urban influence would lead to measurable differences between the two sites however, no statistically significant offset is found between them (see Figure S3).

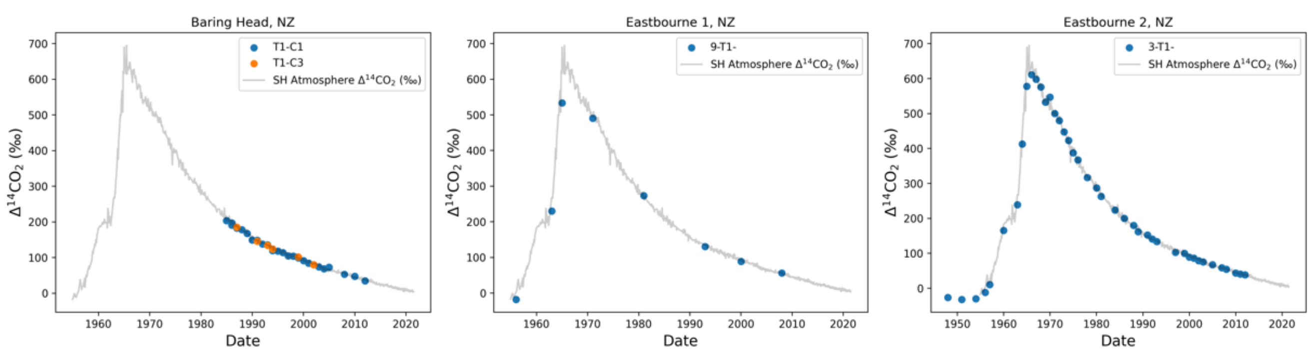

Fig. S7. Baring Head and Eastbourne sites (Turnbull et al., 2017):

No data were removed from any of these sites during ring-count validations. The Baring Head pine is 10m from the 14C sampling station, and on the clifftop exposed to oceanic air. The Eastbourne trees are 15 km from Baring Head, on Wellington Harbor.

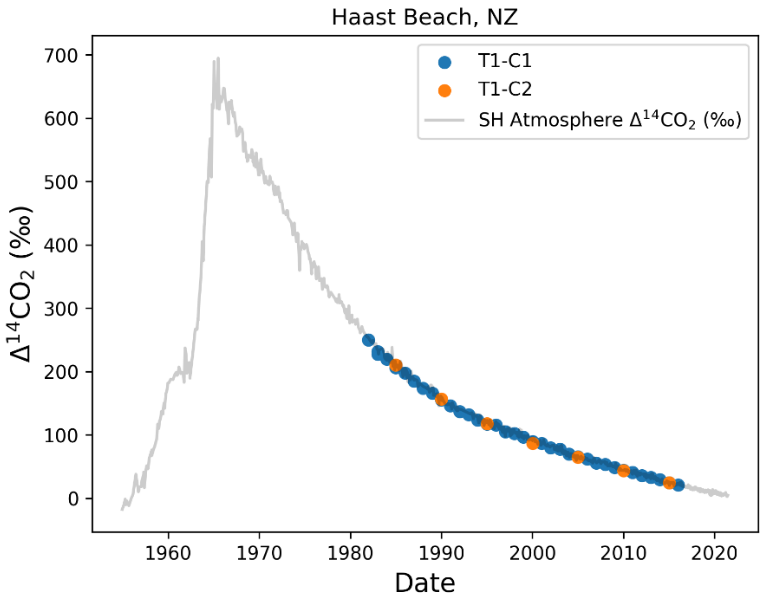

Fig. S8. No data were removed during ring-count validation.

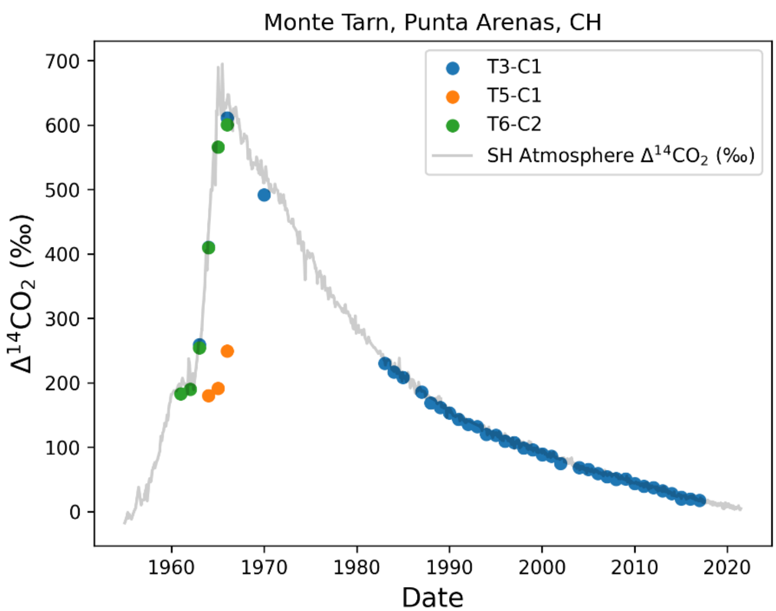

Fig. S9. Tree 5, Core 1 was removed because it does not match the bomb spike

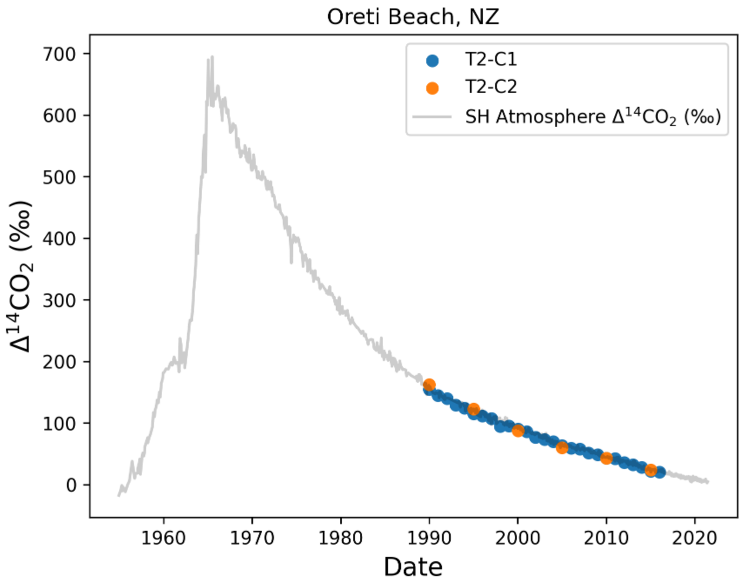

Fig. S10. Cores from Oreti Beach agree as far back as 2000 and then the 1995 and 1990 pairs diverge indicating a ring count error in one or other core. All samples from 1999 back are therefore suspect and are not used.

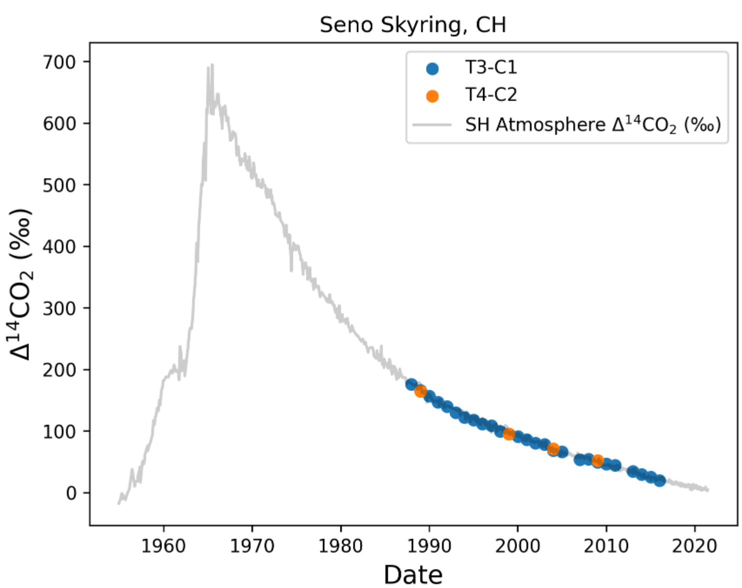

Fig. S11. No data removed from this site

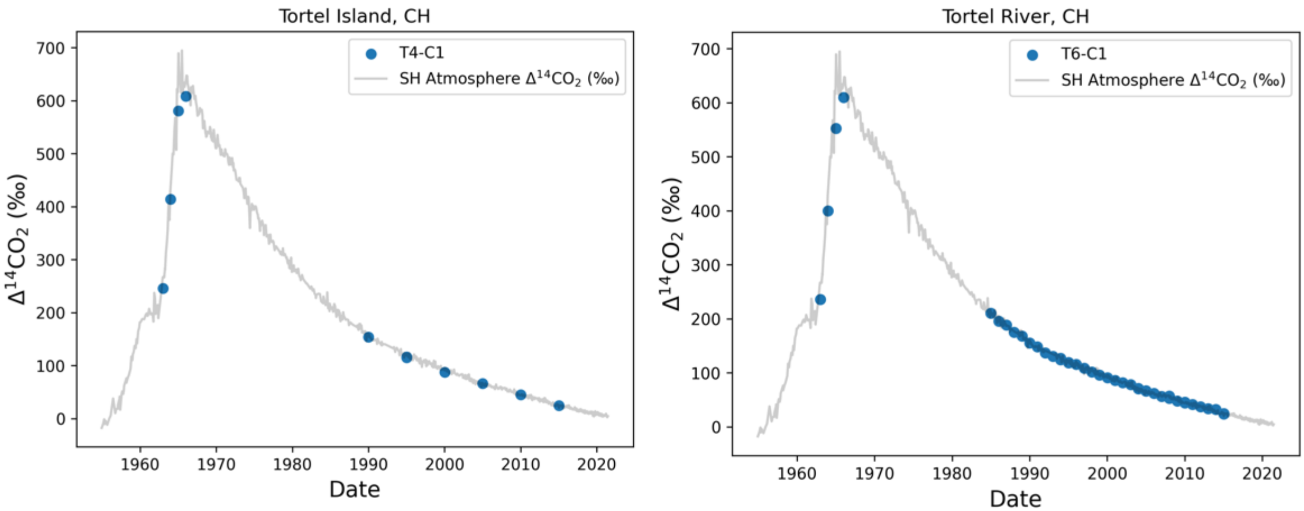

Fig. S12. No data removed from these sites.

S6. Remaining HYSPLIT back-trajectories

Below, all HYSPLIT back-trajectory heat maps are shown. Only a sub-set of 4 sites are shown in the main text.

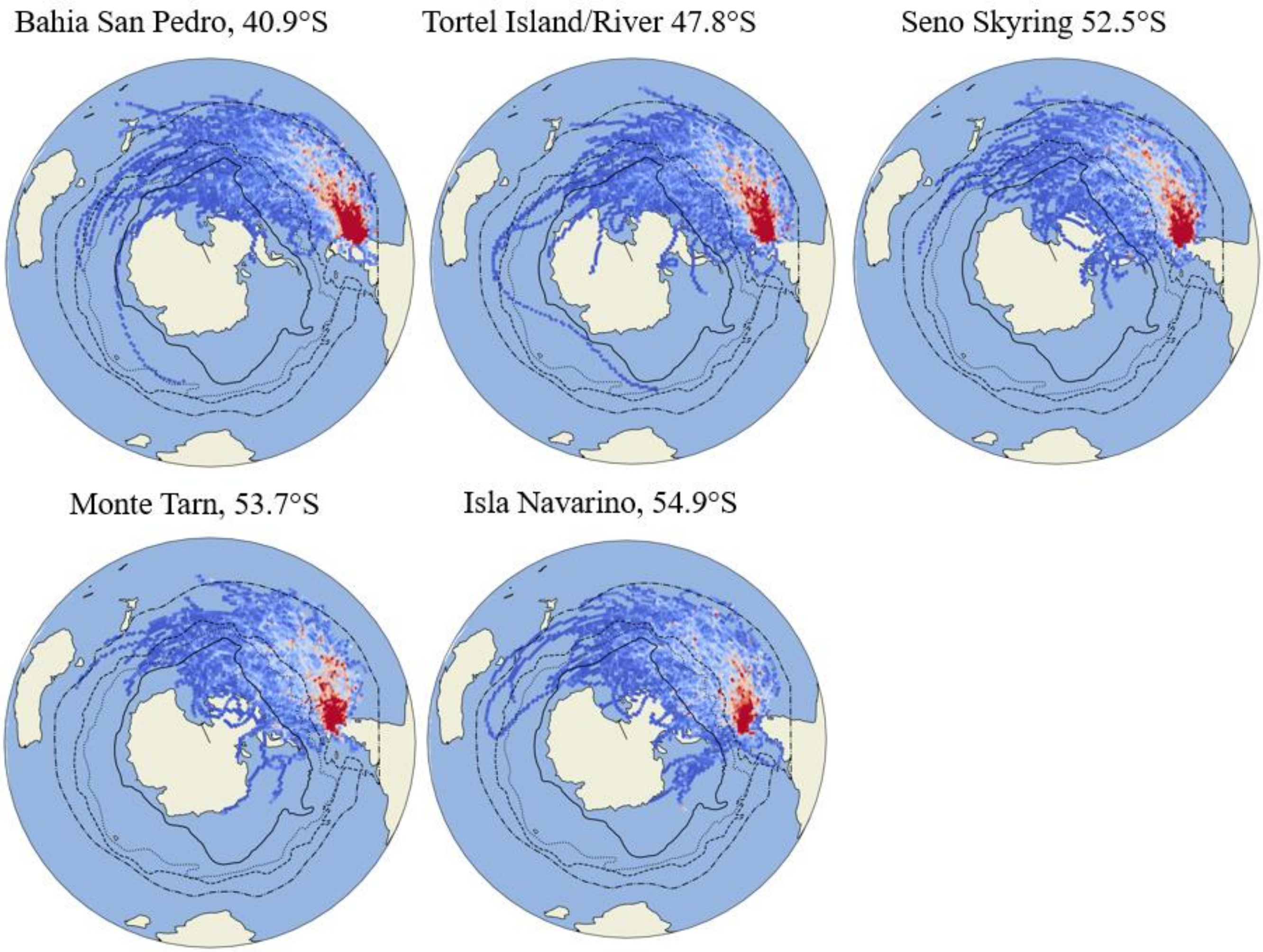

Fig. S13. HYSPLIT back-trajectory heatmaps for sites in Chile.

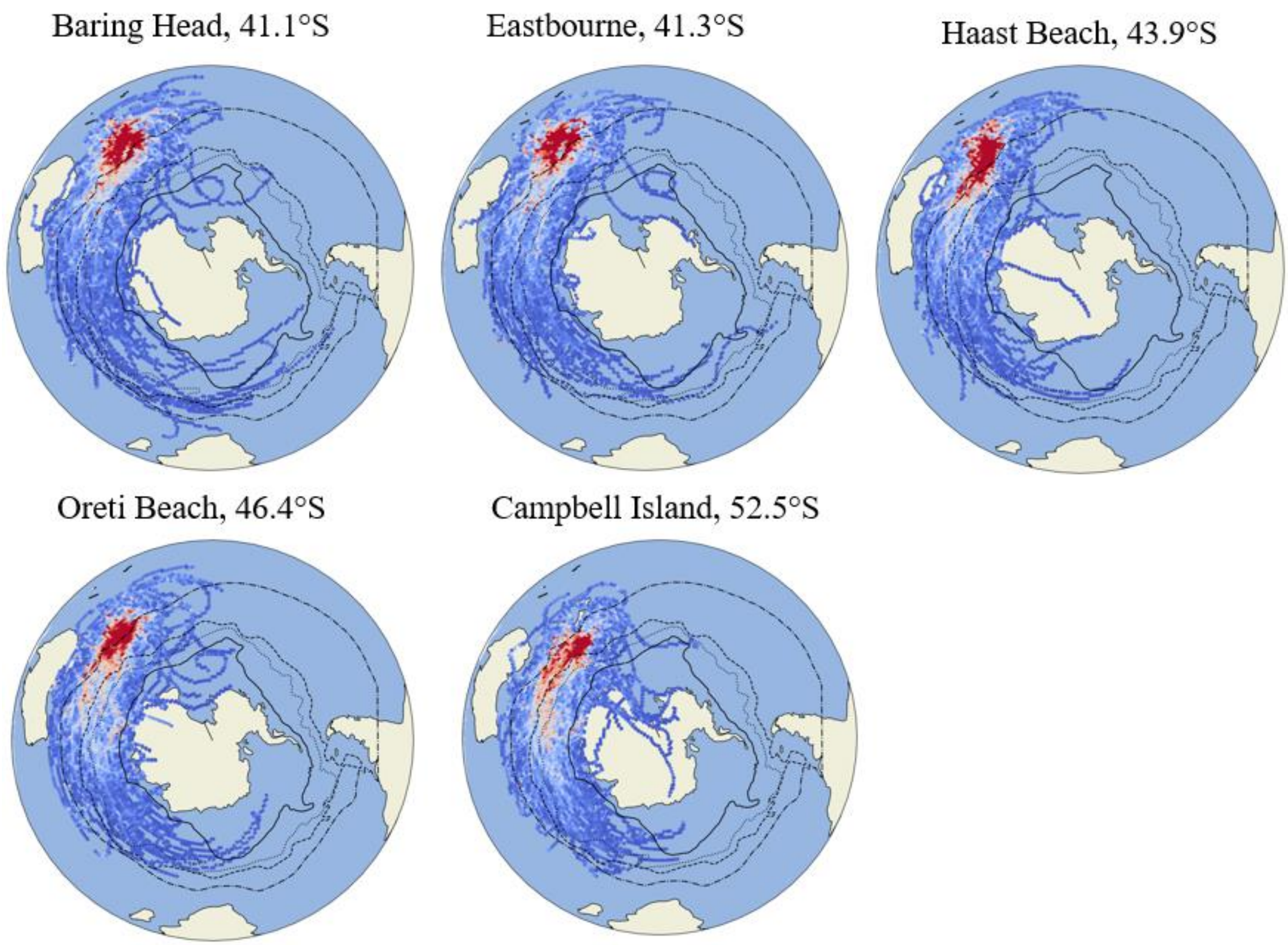

Fig. S14. HYSPLIT back-trajectory heatmaps for sites in New Zealand.