Lund, October 5th, 2024

Carlos Gómez-Ortiz
Department of Physical Geography and Ecosystem Science
Lund University
Sweden

Jens-Uwe Grooß
Editor assigned to Research article EGUSPHERE-2023-2215.
Atmospheric Chemistry and Physics (ACP)

Dear Editor,

Here we address the latest review of our manuscript titled "Can Δ14CO2 observations help atmospheric inversions constrain the fossil CO2 emission budget of Europe?"

We sincerely appreciate feedback on our manuscript and are committed to improve it further.

Below, we provide a detailed response (in regular font) to each of the referee's comments (in italics), indicating how we have addressed them in the revised manuscript. We hope this clarifies any misunderstandings and demonstrates our commitment to meeting the high standards of the journal.

Referee's comments

This manuscript describes the implementation of a dual-tracer approach (CO2 mixing ratios and radiocarbon isotope ratios, Δ14CO2) in an atmospheric inversion framework, LUMIA, for the co-optimization of fossil emissions and land-biosphere fluxes of CO2.

While overall the results appear sound, there are a few points in the description of the methodology that should be clarified before final publication.

Moreover, I am not convinced that this study fully answers the question in the title: "Can Δ14CO2 observations help atmospheric inversions constrain the fossil CO2 emission budget of Europe?" and wonder if a different title would better reflect the scope of the study. The OSSEs carried-out in this study, although fine for testing the implementation of the dual tracer approach, are a bit limited in terms of answering this question. Specifically, the prior error in the OSSE inversions is known, as it is determined from the difference between the true and prior flux datasets. This is not representative of the reality, when the prior error is unknown. Also, the transport in the OSSEs is perfect, which is also not reflecting the reality. Furthermore, to fully answer the question in the title, would entail investigating which of the sampling strategies, that is, i) hourly integrated samples every 3-days, ii) 2-weekly integrated samples, or iii) both, would provide the best constraint on fossil fuel CO2 emissions. This is not to say that these aspects must be covered for the manuscript to be accepted for publication, only that the title should perhaps better reflect the scope of the present study.

We agree with the reviewer that the title might be misleading and suggests results beyond the scope of our methodology and described study here. We propose the new title: “A CO_2 - $\Delta^{14}CO_2$ inversion setup for estimating European fossil CO_2 emissions”.

Specific comments

L14-15: This sentence is unclear, needs more context, do the authors refer to the posterior biosphere fluxes which are retrieved with bias, or something else?

Indeed, we refer to the posterior fluxes. These lines were modified as follows to give more context:

“In all experiments, regions with low sampling coverage, such as Southern Europe and the British Isles, show poorly resolved posterior fossil CO_2 emissions. Although the posterior biosphere fluxes in these regions follow the seasonal patterns of the true fluxes, a significant bias remains, making it impossible to close the total CO_2 budget.”

L46-47: While it is correct that inverse modelling systems that only constrain land-biosphere fluxes assume that the fossil CO2 emissions are well-known, it does not follow that “this is to avoid any bias the fossil CO2 flux might introduce to the estimates of terrestrial fluxes”. Rather the opposite, a fossil CO2 flux estimate that is biased but assumed not to be will introduce errors in the terrestrial fluxes. Furthermore, even in systems constraining only biosphere fluxes, the uncertainty of fossil CO2 emission can be (and should be) accounted for in the observation space.

Indeed, this sentence was misleading and we agree with the reviewer that prescribing wrong fossil CO2 emissions would lead to a bias in the inferred land-biosphere fluxes. Hence, we removed the sentence.

L102: Please change “CO2 concentration” to “CO2 mixing ratios” (or “mole fractions”) as it is the volume mixing ratio (or equivalently mole fraction) that is reported, not the concentration. Please change this elsewhere in the manuscript as well.

We changed the term concentration by mixing ratios in the whole manuscript.

Eq. 1a and 1b: Please use standard notation. In these equations presumably y_co2 and yb_co2 are scalars and Fc is a vector representing 2D space?

Indeed, y and y^b in Eq. 1a and 1b are scalars. We modified the equation and the subsequent mentions in the text.

L150: For completeness please also describe what is y_c_delta_14C. Is this the mixing ratio of 14C-CO2?

Yes, it refers to the mixing ratio of $CO_2 \times \Delta^{14}CO_2$. To add clarity, we reformulated this paragraph as follows:

"where y is the assumed CO_2 and $C\Delta^{14}C$ mixing ratio, yb is the modeled CO_2 and $C\Delta^{14}C$ background mixing ratio (i.e., the boundary condition) (see Section 3.3). Since the values of $\Delta^{14}CO_2$ in ‰ (permil) units are not additive (as it represents the change of the ^{14}C:^{12}C atmospheric ratio relative to an absolute standard of ^{14}C from 1950 (Stuiver and Polach, 1977)), we convert all $\Delta^{14}CO_2$ values to values of $CO_2 \times \Delta^{14}CO_2$ (or $C\Delta^{14}C$ for simplification) (Basu et al., 2016). In terms of units, for mixing ratios this would be $C\Delta^{14}C$ ppm ‰, and for fluxes PgC ‰ yr^{-1}. Since ‰ only means multiplication by 1000, we drop that factor from $\Delta^{14}C$ into the quantity $C\Delta^{14}C$, expressing it in ppm for mole fractions and PgC yr^{-1} for fluxes to maintain the same order of magnitude and units for CO_2 and $C\Delta^{14}C$. For example, a sample with a CO_2 mole fraction of 400 ppm and a $\Delta^{14}C$ value of 45 ‰ would have $C\Delta^{14}C$ = 18 ppm. Expressed in this way, $C\Delta^{14}C$ becomes additive and can be transported by a model. {…}"

L150: Here the authors state that y^b is the "modelled background", whereas in L142, they state that y^b is "calculated by computing a smoothed and detrended average of real observations". Please clarify which of these is it?

We changed the word "modelled" by "assumed" in L150 to make it consistent with the sentence in L142. Since we are doing perfect transport OSSEs, we are using the same background for calculating the synthetic observations and performing the inversions, focusing only in the regional component.

L153: I think in Eq. 1b it should rather be the fraction of 14C in F_c and not the isotopic signature, which represents the ratio of 14C in the sample relative to the reference, and y_c_delta_14C would be the mixing ratio of 14C-CO2.

We agree with the referee. We modified these lines as follows:

"In Eq. 1b, the term $\boldsymbol{\Delta}_c$ represents the fraction of ^{14}C in the accompanying flux category $\boldsymbol{F}_c$ (Tans et al., 1979; Turnbull et al., 2016)."

L155: Similar to the above comment, to calculate the mixing ratio of 14C-CO2 one would need to multiply CO2 mixing ratio by the fraction of 14C-CO2, not delta_14C. Or unless the authors use the assumption that 14C << 12C and thus the ratio 14C/12C is approximately equal to 14C/(12C + 14C) in which case this should be explicitly stated.

We modified this paragraph to add clarity as answered above. We use $\Delta^{14}CO_2$ as defined by Stuiver & Polach (1977), since this is how ICOS samples are reported. An approximation of this definition is:

$$\Delta^{14}C(‰) = \left(\frac{^{14}C/^{12}C_{sample}}{^{14}C/^{12}C_{standard}} - 1\right) \times 1000$$

L155: "ppm" is a unit of mixing ratio not concentration.

We changed the term concentration by mixing ratios in the whole manuscript.

*L157: Again, if the fraction of 14C/12C is used rather that delta_14C, which I think it should be, then the units of delta_c*F_c will be PgC/yr. The unit of PgC_permil/yr does not correspond with y_c_delta_14C, which is ppm.*

Indeed, we modeled $C\Delta^{14}C$ in units of ppm for mixing ratios and e.g., PgC yr^{-1} for fluxes. We have modified this paragraph as answered above.

L170: Since fossil CO2 does not contain any 14C it does not contribute to a change in the mixing ratio of 14C-CO2, i.e., has no effect on y_c_delta_14C (in Eq.1b).

We respectfully disagree with the referee. Although fossil CO_2 does not contain any ^{14}C, it does contribute to a change in the $\Delta^{14}C$ of atmospheric CO_2 by diluting the amount of ^{14}C in the atmosphere. This dilution leads to a reduction of the $C\Delta^{14}C$ mixing ratio and, consequently, the $\Delta^{14}CO_2$ isotopic ratio. This process is the basis of the Suess effect (Suess 1955; Tans, De Jong, and Mook 1979), and it is the fundamental reason for using $\Delta^{14}CO_2$ as a tracer to separate the fossil and the natural components in atmospheric CO_2 observations (Turnbull et al. 2009; Turnbull, Graven, and Krakauer 2016).

L204: Why was the 2-week integrated sampling strategy for delta_14C chosen, rather than the 1-hour integrated sample every 3-days? Surely, the 1-hour samples would better help resolve the fossil fuel signal, since the transport and source regions could change significantly over the course of 2 weeks.

The ICOS Atmosphere network has been collecting 2-week integrated samples of $\Delta^{14}CO_2$ since 2016 at 12 stations across Europe thus it is important to us to evaluate the potential use of the available data. Nevertheless, we are aware of the limitations of the 2-week samples and we recently submitted a new manuscript to ACP exploring different sampling strategies.

L240: Instead of "grid points" do the authors rather mean "grid cells"?

We changed this line to "grid cells".

*Eq.7: What is the matrix operation indicated by ⊗ ? I read it to mean the Kronecker product, in which case T_H ⊗ T_T would have dimensions (n^p_mod*n^t_mod, n^p_opt*n^p_mod), and then x_c would need to be a vector of n^p_opt*n^p_mod. Please confirm if this is correct? It would help if the dimensions of H and x_c were also given.*

The reviewer is correct that this is a Kronecker product, but it results is a $\left(n_{mod}^{p} * n_{mod}^{t}, n_{opt}^{p} * n_{opt}^{t}\right)$ matrix. We have added the dimensions of $\boldsymbol{x}_c$ and $\mathbf{H}$ to the sentence following the equation, to lift any source of doubt on the reader side:

"where $\mathbf{H}$ is the observation operator with dimensions $\left(n_{obs}, n_{p_{opt}} * n_{t_{opt}}\right)$, and $\boldsymbol{x}_c$ with dimensions $\left(n_{p_{opt}}, n_{t_{opt}}\right)$ is the portion of the control vector $\boldsymbol{x}$ that contains offsets for the optimized categories c. The matrices T_T $\left(n_{t_{mod}}, n_{t_{opt}}\right)$ and T_H $\left(n_{p_{mod}}, n_{p_{opt}}\right)$ contain the relative contribution of each model time step t_{mod} (1 hour) and of each grid cell p_{mod} (0.5° × 0.5°) to each optimized time step t_{opt} and cluster p_{mod}, with $n_{t_{opt}}$ and $n_{p_{opt}}$ the number of optimized intervals (weekly) and grid cell clusters, respectively."

L259: The definitions of L_h and L_t should be included here.

We added the following sentence to the end of this line:

"L_h and L_t represent the horizontal and temporal correlation lengths, respectively."

L381-382: The time window over which the standard deviation is calculated (7-days), which is used as a proxy for the observation error, is very long. This would imply that the authors do not have much confidence in the model's ability to represent synoptic variability in the mixing ratios. There is no discussion of why this long time window was chosen or the evaluation of this choice, e.g., how well does the model capture the variability of tracers for which the fluxes are likely better known (Radon or SF6)?

This is a valid point, and to test this we repeated the ZBASE experiment with two additional time windows: half a week, and one day. We found the standard deviation to not be too sensitive to the window width (in average at each site) and also, there is no significant impact on the posterior results (Fig. 5 and 6 of this document). In inversions against real observations, we would fine tune the observation uncertainties based in part of the quality of the prior fit to the data, but we cannot do this here since we do not assimilate real data. Furthermore, it is common in inversions to "inflate" the uncertainties, to compensate for the fact that the observation uncertainties are treated as independent (i.e. the "R" matrix is diagonal), which isn't accurate. Our observation error values are on the same order of magnitude as what is typically used in LUMIA CO2 inversions. The table below reports for instance the values used in Munassar et al. (2023). Note that these are weekly aggregated uncertainties. For comparison we have calculated (average) weekly aggregated uncertainties in our case.

Site	**Averaged observation errors per site (ppm)**				
	1 week	**Half week**	**1 day**	**Weekly uncertainty Munassar et al.**	**Weekly uncertainty as Munassar et al.**
BIR	6.76	6.37	5.62	2.5	1.1
CMN	5.9	5.58	4.96	1.5	1.0
GAT	13.27	12.38	10.87	1.5	2.2
HPB	12.19	11.35	10.06	1.5	2.1
HTM	11.8	11.06	9.74	1.5	2.0
IPR	19.94	18.97	16.72	1.5	3.4
JFJ	4.63	4.36	3.9	1.5	0.8
KRE	12	11.19	9.73	1.5	2.0
LIN	18.09	16.8	14.95	2.5	3.1
LMP	5.09	4.83	4.35	1.5	0.9
LUT	17.28	16.16	14.18	2.5	2.9

NOR	10.03	9.34	8.19	1.5	1.7
OPE	15.82	14.65	12.94	1.5	2.7
PAL	6.78	6.26	5.57	2.5	1.1
PUI	5.77	5.47	4.82	1.5	1.0
PUY	7.31	6.83	6.12	1.5	1.2
RGL	7.87	7.32	6.59	1.5	1.3
SAC	27.9	25.88	23.27	2.5	4.7
SMR	9.14	8.51	7.52	1.5	1.5
SSL	7.16	6.78	6.06	1.5	1.2
SVB	7.65	7.12	6.41	1.5	1.3
TRN	14.66	13.52	12.03	1.5	2.5
UTO	10.52	9.91	8.85	1.5	1.8
WAO	14.73	13.62	12.26	1.5	2.5

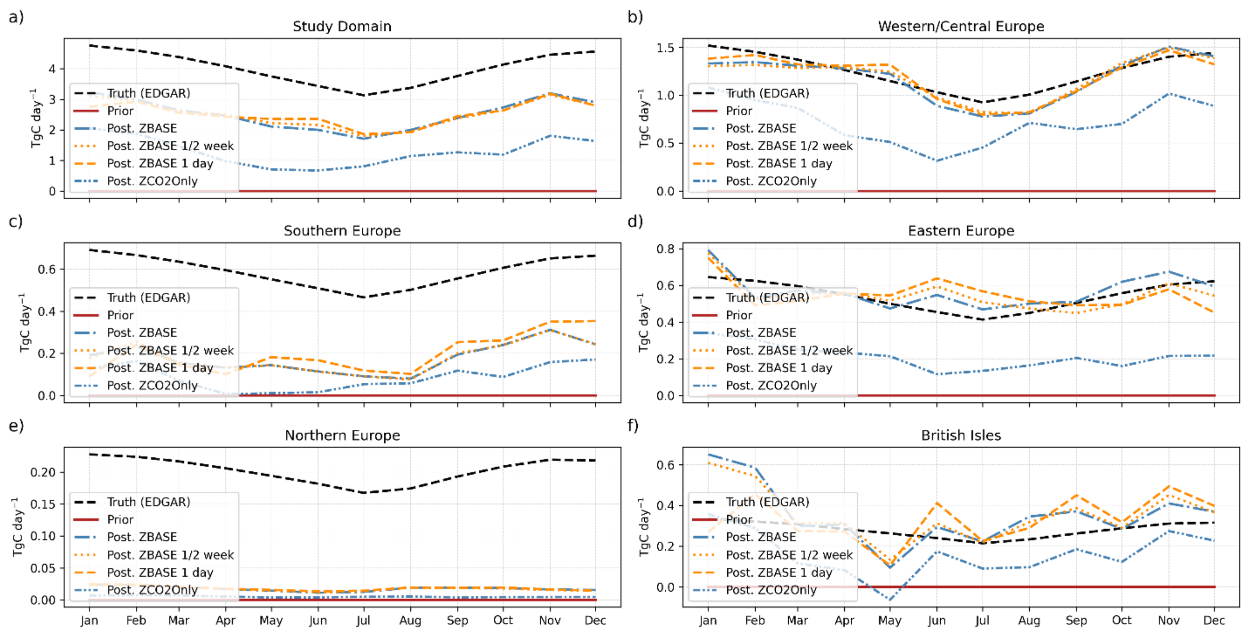

Figure 5. Fossil CO_2 emissions.

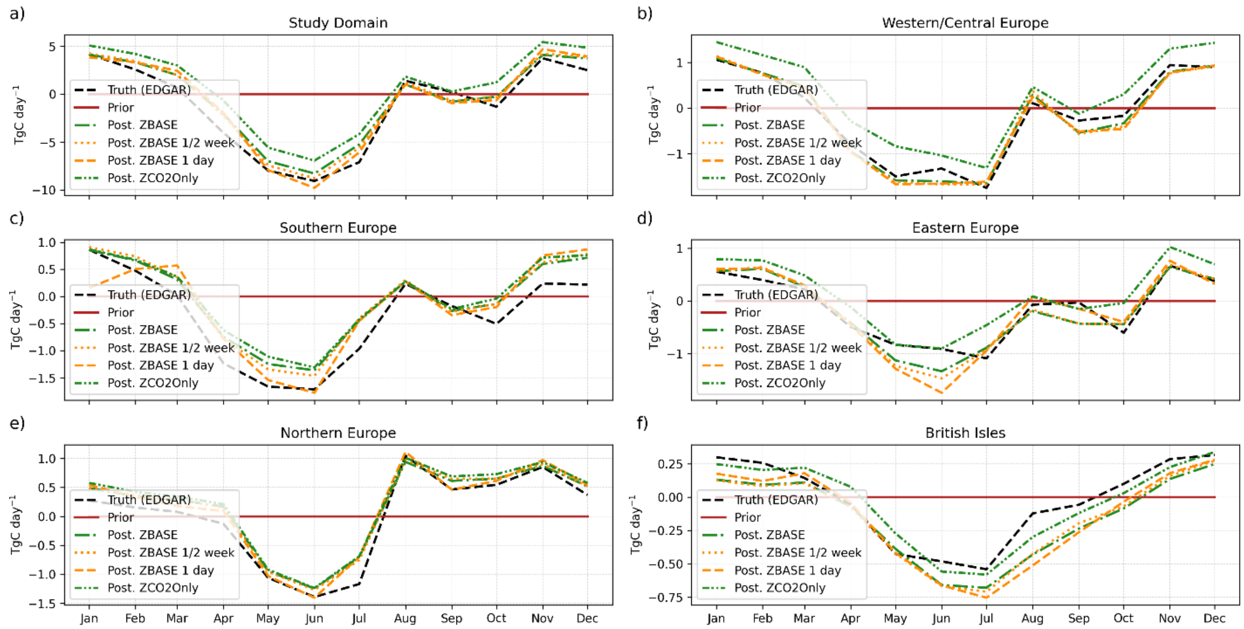

Figure 6. Biosphere (NEE) fluxes.

L392-393: I think the authors should specify that the prior fluxes for F_ff and F_bio can have similar distributions to the "true" fluxes, otherwise it's not clear if by "similar" the authors mean similar to each other or what similar to what?

We reformulate this sentence as follows:

"The reason for using prior fluxes set to zero is that the flux products for both categories can have spatial and temporal distributions similar to their respective true values, making it easier for the model to retrieve the true fluxes."

L402-403: I do not follow how the potentially large error in F_biodis can be accounted for by using the true value in the inversion and not optimizing it?

We modified the description of the experiment as follows to add clarity:

"In the final inversion, BASENoBD, we prescribe F_{biodis} (i.e., the true value in this context) instead of optimizing it. The terrestrial disequilibrium term (F_{biodis}) is challenging to estimate due to the large uncertainties associated with heterotrophic respiration fluxes and the age of respired carbon (Basu et al., 2016). These uncertainties can vary significantly depending on the vegetation model or methodology used. We compare the posterior F_{ff} of this experiment with the one of the BASE experiment (in which F_{biodis} is optimized), to evaluate the impact of the prior F_{biodis} product on the posterior F_{ff}. By keeping F_{biodis} fixed in BASENoBD, we can assess how much of the error in the posterior F_{ff} of BASE comes from the additional optimization of F_{biodis}."

L425: It is not clear what is being compared here, the ZBASE and ZCO2ONLY inversions are closer in agreement to the truth compared to what? The prior?

The sentence is intended to highlight the comparison between the agreement of the truth and the posterior for F_{bio} relative to F_{ff} in the ZBASE and ZCO2Only experiments. To clarify this, we revised the sentence as follows:

"In general, there is a closer agreement between the posterior and the truth for the biosphere fluxes (F_{bio}) than for the fossil CO_2 emissions (F_{ff}) in both the ZBASE and ZCO2Only experiments. This means that the model performs better at recovering F_{bio} from the observations compared to F_{ff}, as shown in Figure 6 for F_{bio} and Figure 5 for F_{ff}."

L428: "ZBASE exhibits a closer alignment to the posterior" – do the authors rather mean that the posterior of ZBASE agrees better with the truth?

Yes, we mean that the posterior F_{ff} of ZBASE agrees better with the truth than the one of ZCO2Only. We modified this sentence as follows:

"Specifically, the posterior F_{ff} ZBASE exhibits closer alignment to the truth than ZCO2Only with a lower RMSE (see Table 4), indicating a better fit of the seasonality for F_{ff}."

Figure 9: I think the authors should discuss why in July (especially in Eastern Europe) there is this strong departure from the prior and from the true emissions? What is driving this?

We do discuss this in L630-L637. We added at the end of these lines a new sentence (highlighted):

"As shown in Figure 11, the maximum difference between the prior and the true Fbiodis is of the same order of magnitude for Western/Central Europe (2.1 TgC day^{-1}) and Eastern Europe (1.3 TgC day−1) in July. For F_{ff}, however, the difference between the prior and truth is about one order of magnitude larger for Western/Central Europe compared to Eastern Europe (0.03 vs 0.005 TgC day^{-1}). This larger difference causes a stronger dilution of the fossil emissions in Eastern Europe, and therefore essentially lowers the signal-to-noise ratio of the $\Delta^{14}CO_2$ measurements, and added to the lower network coverage compared to Western/Central Europe, a poorer constrain of the fossil CO_2 emissions. As seen also in Figure 9, this is particularly evident in Eastern Europe during the summer months, where the fossil CO_2 signal is further convoluted by the large biospheric uptake, making it more difficult to accurately constrain fossil emissions in this region."

Discussion:

How do the diurnal cycles of biosphere CO2 fluxes differ between LPJ-GUESS and VPRM? The diurnal cycle is not optimized in LUMIA (weekly means only are optimized) thus I was wondering how sensitive is the inversion to differences in the diurnal cycle – or are the uncertainties in the observation space so large that this does not have much of an impact?

Thank you for raising this important point. We agree that the diurnal cycles of biospheric CO_2 fluxes are an interesting aspect to explore, particularly when comparing LPJ-GUESS and VPRM. However, in the context of our LUMIA implementation, we focus on weekly means, and the diurnal cycle is not directly optimized in the inversion process. The foreground part of the observations is sensitive to fluxes aggregated over a few days, which naturally attenuates the impact of diurnal cycles in the observations.

Moreover, since we use afternoon-only data, the inversion is not designed to resolve the full daily cycle of CO_2 fluxes. While differences in the diurnal cycles between LPJ-GUESS and VPRM might exist, we don't expect these differences to significantly impact the results of the inversion. The uncertainties in the observation space, combined with the aggregation over days, likely minimize the sensitivity to variations in the diurnal cycle.

We also consider that this aspect is slightly beyond the scope of this paper. Testing this in a pure CO_2 inversion might provide valuable insights. However, in practice, we rely on the daily cycles provided by vegetation models, as the "true" daily cycle remains uncertain. Thus, optimizing the diurnal cycle is not a primary focus of the inversion setup at this stage.

We appreciate your insightful question and think it could be a great direction for future research in a more dedicated inversion setup.

Technical comments

L603: should be: "constrained in their inversions" (its -> their)

We fixed this in the text.

References

Munassar, Saqr, Guillaume Monteil, Marko Scholze, Ute Karstens, Christian Rödenbeck, Frank-Thomas Koch, Kai U. Totsche, and Christoph Gerbig. 2023. 'Why Do Inverse Models Disagree? A Case Study with Two European CO_2 Inversions'. *Atmospheric Chemistry and Physics* 23(4):2813–28. doi: 10.5194/acp-23-2813-2023.

Stuiver, Minze, and Henry A. Polach. 1977. 'Discussion Reporting of ^{14}C Data'. *Radiocarbon* 19(3):355–63. doi: 10.1017/S0033822200003672.

Suess, Hans E. 1955. 'Radiocarbon Concentration in Modern Wood'. *Science* 122(3166):415–17. doi: 10.1126/science.122.3166.415.b.

Tans, P. P., A. F. M. De Jong, and W. G. Mook. 1979. 'Natural Atmospheric 14C Variation and the Suess Effect'. *Nature* 280(5725):826–28. doi: 10.1038/280826a0.

Turnbull, J. C., H. Graven, and N. Y. Krakauer. 2016. 'Radiocarbon in the Atmosphere'. Pp. 83–137 in *Radiocarbon and Climate Change: Mechanisms, Applications and Laboratory Techniques*, edited by E. A. G. Schuur, E. Druffel, and S. E. Trumbore. Cham: Springer International Publishing.

Turnbull, Jocelyn, Peter Rayner, John Miller, Tobias Naegler, Philippe Ciais, and Anne Cozic. 2009. 'On the Use of 14CO2 as a Tracer for Fossil Fuel CO2: Quantifying Uncertainties Using an Atmospheric Transport Model'. *Journal of Geophysical Research: Atmospheres* 114(D22). doi: https://doi.org/10.1029/2009JD012308.