Lund, July 19th, 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. But in this round of review, we are partly facing again questions that have already been answered and supported by relevant bibliography in previous rounds as well as new concerns that have not been mentioned before. While we have made modifications based on the latest comments to clarify the manuscript for the referee, we consider these new comments rather to be minor revisions. However, we are concerned that this process may become an endless loop of revisions, as there seems to be a fundamental misunderstanding of the concept of our study by the referee. This may be caused by preconceived conceptions of their field of research from the assessment of our manuscript (examples for this are the confusion about the grid, the usage of the term 'Observation Simulation System Experiment', and the comment on basic linear algebra as explained in more detail below). We have also noticed that the referee has progressively lowered the rating of the manuscript in terms of scientific significance, scientific quality, and presentation quality, despite acknowledging improvements made in each round of review. Therefore, we suggest that a direct contact with the referee may help to clarify their concerns.

We are puzzled about the referee's question on basic linear algebra concepts, such as "However, it is unclear how B^-1 and x multiply in (6),". As mentioned above we assume that this is caused by a fundamental misunderstanding of our manuscript. This misunderstanding appears to arise from an earlier stage of the manuscript, leading to questions about basic concepts. We cannot adequately address the referee's concerns without knowing the specific source of this confusion. It seems the referee believes our study involves a more complex methodology than it actually does. Our manuscript describes a classical sensitivity experiment in a variational inverse modeling application, using standard optimization algorithms common in the field and a general methodology described in many publications and also more specifically already elsewhere (e.g. Monteil and Scholze, 2021). This is not a model development manuscript and does not employ any unique optimization techniques.

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

The updated manuscript contains a number of improvements. However, I also see a number of points where the authors are reluctant to follow my recommendations.
I summarise the main points below:
comments on the original communication to the editor:
1. "…but it is fundamentally an arbitrary choice, similar to the definition of the domain extent or the selection of a specific resolution for the transport model. As such, it is not feasible to provide a mathematical description for this more than the actual code referenced in the answer below."
As such, it is not feasible to provide a mathematical description for it beyond the actual code referenced in the response below".
The grid design choices made by the authors may well be reasonable, but they are not arbitrary per se, as the model is expected to produce valid results consistent with the objective of the paper.

Our sentence in the rebuttal directly addressed the referee's request for a "mathematical foundation for Appendix B" (as noted in the last line of their previous review). Appendix B outlines the procedure for defining the reduced grid on which the control vector is based. It is important to note that this is distinct from the transport model grid, which was mentioned in the manuscript (lines 241-245). The choice for the control vector grid is indeed arbitrary, based on the authors' best judgment rather than a formal derivation. We justified this choice in lines 234-240 of the manuscript, a justification to which the referee has not previously objected.

I understand that a formal study of model optimisation is beyond the scope of the paper. However, scientific scrutiny deserves at least that the model grid design (extension, resolution) and assumed emission patterns are based on a discussion of typical advection timescales and minimum error estimates for emission source assumptions, summarising the underlying assumptions and assuring the reader of the limits of possible modelling errors. I do not understand why this moderate request for more concern and awareness of possible design problems is apparently seen as an imposition by the authors. One sentence should suffice.

The referee's comment suggests to us that they misunderstand our grid design. All previous questions were about the optimization grid shown in Figure 2, which is not at all related to advection timescales. Our transport calculations are done at a 0.5° spatial and hourly temporal resolution, and emissions are provided at the same resolution. The optimization grid is used only to solve the inverse problem (i.e., the control vector), and advection timescales are not relevant.

This new request for justification of resolution and domain definition was not raised before. The domain choice has been commented since the first version of the manuscript at the beginning of Section 2.1. Regional transport model (Section 2.2. Observations, L189-190 of the last version) and supported by publications such as Monteil et al., 2020; and Thompson et al., 2020 (cited on the manuscript) and other ACP and Copernicus publications such as McGrath et al., 2023; Munassar et al., 2023; and Petrescu et al., 2021. We have already explained and justified our

choices for the optimization grid in the manuscript. We think that our current explanation is sufficient and appropriate for the scope of this manuscript (as it was in the previous publications mentioned before).

2. The algorithmic description is still a mystery to me. In some exemplary detail: The authors have (L222): "H is the observation operator, which includes the transport model…". Since, according to eq. (6), H (including the time propagating transport model) acts on the vector x to be optimised, the latter presumably according to the Ide formulation.

Indeed, our formulation was somewhat confusing and now we have changed our equations 1 to 7 to account for that. In equations 1-4 we have introduced the regional transport model operator K. Equations 5-7 have been reformulated to clarify the linearity of the observation operator H.

3. The H M(tangent linear model)(t_i, t_0) acts on x(t=0), which is not a multi-time step state vector but rather an observation operator H. The background term, with the inverse space-time covariance matrix B (as previously stated in 2.3.2), would be in compliance with this if the generalized observation operator H were also applied. However, this is absent, and there is no indication of how the time propagation is managed for this 1. rhs term of (6).

We adjusted the cost function (Equation (6)) to write it in the common form used by other authors, and removed the sentence in L219-220 of the latest version "that are available over a range of times, also known assimilation window" which might be misunderstood as that we do sequential assimilation, but otherwise, we do not understand the referee's concern. We do not really understand the comment of the reviewer. There is no tangent linear model in our case, since $\mathbf{H}$ is linear. $\mathbf{H}$ acts on $\boldsymbol{x}$ in its entirety, not on $\boldsymbol{x}(t)$. It's not clear to us what the referee means by "time propagation": there is a time propagation within $\mathbf{H}$ (i.e. the observations at time $t = t1$ are influenced by control variables $\boldsymbol{x}(t \leq t1)$, but there is no time propagation of the solution (i.e. $\boldsymbol{x}(t = t1)$ doesn't depend on $\boldsymbol{x}(t < t1)$).

Furthermore, the presentation of the covariance formulation is not sufficiently clear.
L 265: cov(x1, x2) = σx1 σx2 exp(−(d(p1,p2)/Lh)2exp(−|t2−t1|/Lt)
This expression should be indexed properly, that is, with time and location indices running independently, and both by different index variables with associated limits. This would ensure that the reader can identify the entry-wise construction and associated x.

The equation is strictly correct, and exactly how it's used in the code. The covariance between two elements $\boldsymbol{x}_1$ and $\boldsymbol{x}_1$ of the control vector is the product of their variances ($\sigma\boldsymbol{x}_1$ and $\sigma\boldsymbol{x}_2$), and of two decorrelation functions, decreasing exponentially with the spatial and temporal distance between the two points. There is zero ambiguity in the Equation. Nevertheless, we replaced the sub-indexes 1 and 2 in the equation by i and j to put it in a more general way. The indices refer neither to time nor to location, they refer to the position of the variables in the control vector. x_i has coordinates (p_i, t_i), and x_j has coordinates (p_j, t_j).

The Equation at L265 gives exactly the entry-wise construction of the covariance matrix. The only step missing in this equation is for the case when x_i and x_j belong to different emission categories, in which case the correlation is 0. But this is very clearly specified just above, at L252: "we assume no correlations between different categories and different tracers".

However, it is unclear how B^-1 and x multiply in (6).

1. $\boldsymbol{x}$ has $n = n_t^{opt} \times n_p^{opt} \times n_{cat}$
2. $\mathbf{B}$ has (n, n) elements
3. The product $\boldsymbol{x}^T \mathbf{B}^{-1}$ results in a n dimension vector

Section 2.3.2 (Covariance construction-related issues)
In their response, the authors assert that: "It is our contention that this section provides a lucid, step-by-step explication. In particular, the entries of the prior error covariance matrix B are defined by the equation presented in Section 2.3.2 (L251 of the revised manuscript, originally L204). I must respectfully disagree. I endeavoured to comprehend the material, but upon reaching a point where I believed I had succeeded, I encountered inconsistencies in another location. It is unclear why a section is not devoted to defining the subspace section of vector x (sub)segment wise.

We are puzzled why the referee is misquoting us here. Their paraphrasing suggests a very different tone than we actually used in our previous response. The correct quote is: “We believe this section provides a clear, step-by-step explanation. Specifically, the entries of the prior error covariance matrix $\mathbf{B}$ are described by the equation found in Section 2.3.2 (L251 of the revised manuscript, originally L204).”

The vector $\boldsymbol{x}$ can be divided in one section for each tracer/category. The approach for each of these sections is the same and is described in L253 ("the sections of $\mathbf{B}$ specific to each tracer/category ...") to L261. The part specific to the subsections is described just after, from L264 to L270. There is no ambiguity in our description of the covariance matrix.

Firstly, it is assumed that the term "offset" refers to what is commonly referred to in data assimilation as an "increment," which is likely an analysis increment rather than an observation increment.

No. The term "offset" refers to the fact that the control vector contains "offsets" to the prior emissions, as opposed to e.g. scaling factors, which are another commonly encountered approach. It refers to the physical quantities that the control vector represents, whereas “increment” refers to the change of the values of the control vector between two iterations of the inversion (regardless of what physical values this “increment” corresponds to).

The term “analysis” used by the referee is commonly used in the context of sequential inversions, such as those employed for weather forecasting, which estimate the atmospheric state through a series of forecast and analysis steps. However, this term is not adapted to atmospheric flux inversions, where the terms “prior” and “posterior” are to be used instead. Beyond the naming convention, this also highlights that there is no time propagation of the solution.

L250: : "The matrix Xc x is the portion of the control vector x that contains offsets for the category c, reshaped as a (ntopt, npopt) matrix, with ntopt and npopt the number of optimized (weekly) intervals and grid-cell clusters, respectively."
In what sense may a matrix X, comprising portions of a vector, be defined in terms of vector calculus? It is possible that this is the source of the misunderstanding.

- $\mathbf{T}_T$ is a (n_t^{opt}, n_p^{opt}) matrix.
- $\boldsymbol{x}_c$ is a $(n_t^{opt} \times n_p^{opt})$ vector, reshaped as a (n_t^{opt}, n_p^{opt}) matrix $\mathbf{X}_c$.
- $\mathbf{T}_\mathrm{H}$ is a (n_t^{opt}, n_p^{opt}) matrix.

The product $d\mathbf{F}_c = \mathbf{T}_T \mathbf{X}_c \mathbf{T}_H$ gives a (n_t^{opt}, n_p^{opt}) matrix (i.e. 1 year of hourly emission offsets at a 0.5° resolution, for one category c). The corresponding emissions are just $\mathbf{F}_c = \mathbf{F}_c^0 + d\mathbf{F}_c$, with $\mathbf{F}_c^0$ the prior emissions for that category (still at a 0.5° and hourly resolution).

The adjoint is simply $\mathbf{X}_c^{adj} = \mathbf{T}_H^T d\mathbf{F}_c^{adj} \mathbf{T}_T^T$, which aggregates a (n_t^{opt}, n_p^{opt}) adjoint emission array (for category c) onto a (n_t^{mod}, n_p^{mod}). The $\mathbf{T}_T$ and $\mathbf{T}_\mathrm{H}$ matrices are category specific. These are basic aggregation/rebinning operations.

The rationale behind these equations is:

- The transport is computed hourly (with hourly emissions), but the inversions solves for weekly offsets (i.e. the emissions within a week will all be increased or decreased by the same absolute amount). An operator is needed to project these weekly offsets onto the hourly emissions.
- Likewise, in the spatial dimension, the offsets are defined on a reduced grid (Figure 2), whereas the emissions themselves are defined on a 0.5° resolution, the aforementioned operator also needs to be extended to project the emissions from the reduced grid to the model grid.

We have modified Eq. 7 and the section related to it such that we do not use $\mathbf{X}$ in the description. We think that our revised version is easier to follow.

While for partial differential equations, discretized physical domains are transferred to vectors and back to harness vector calculus, the rationale behind the procedure addressed here is unclear.

There are no partial differential equations involved in this procedure, neither are they mentioned in the text or the equations. The rationale is just that at the transport model resolution, there are 42 million spatiotemporal grid cells (24 hours * 365 days * 80 lon * 60 lat), for each emission category, which drops to ~130000 per category with our reduced grid, which is computationally more favorable. Also, the density of the observation network is in any case way too low to robustly resolve the emissions at the model resolution.

Note, once again, that this reduced resolution concerns only the offsets optimized by the inversions. The base emissions, onto which these offsets are added, are still provided and transported at high resolution.

In addition to the ACP paper writing guidelines, which have been partly replicated in my previous review, it is important to ensure that mathematical formulae, symbols, abbreviations and units are correctly defined and used. • Are the scientific methods and assumptions valid and clearly outlined? It is not possible to discern an enhanced derivation from the generic variational formula (6) in comparison to eq. (7), and furthermore, the verbal description of the covariance matrix B in 2.3.2 is not sufficiently clear.

We answer this in previous questions above. Regarding the verbal description of $\mathbf{B}$, we disagree. The description is perfectly clear and unambiguous. As mentioned before, we suspect that there is a more fundamental misunderstanding, which lead the referee to misinterpret this section. We don't know precisely how to solve this misunderstanding without knowing its origin, but it is not in this section.

Please refer to Section 3. The confusion surrounding the use of the term "OSSE" and its relation to identical twin (IT) experiments
Sect 3. The confusion surrounding the use of the acronym "OSSE" in relation to identical twin (IT) experiments: Although the authors acknowledge that their approach is correctly classified as an IT, their justification by customary use in their atmospheric inversion community is not valid. The authors might not make false use of a terminology with precise distinctions between (OSSE – IT) from data assimilation without any necessity, neglecting precision.
The authors may not be employing the correct terminology with regard to the precise distinctions between OSSE and IT in the context of data assimilation, without any necessity for such precision.

We have already addressed this question in the past and we are surprised by the invalidation of our terminology, especially given the established practices in our research community. Numerous papers within our community (also published in ACP and already cited in previous responses) support the usage of the terminology we have employed. It is important to recognize that different research communities may develop distinct terminologies and methodologies, even if the underlying principles are similar. This divergence in terminology can cause confusion to the referee, but it is consistent with the practices in our field of atmospheric inverse modeling.

Our intention was not to neglect precision but to adhere to the accepted conventions within our community. We hope this explanation clarifies our approach and the reasoning behind our terminology.

We have also clearly stated in the manuscript that we are performing a ‘perfect transport Observing System Simulation Experiment’ (L 282).

Note: Given the questions raised by the referee about the rigorousness of the mathematical formulations, in the attached document we provide a detailed derivation of all the equations in the methodological section of the manuscript. However, we consider such level of detail is too detailed for the paper.

References

McGrath, M. J., Petrescu, A. M. R., Peylin, P., Andrew, R. M., Matthews, B., Dentener, F., Balkovič, J., Bastrikov, V., Becker, M., Broquet, G., Ciais, P., Fortems-Cheiney, A., Ganzenmüller, R., Grassi, G., Harris, I., Jones, M., Knauer, J., Kuhnert, M., Monteil, G., … Walther, S. (2023). The consolidated European synthesis of CO2 emissions and removals for the European Union and United Kingdom: 1990-2020. *Earth System Science Data*, *15*(10), 4295–4370. https://doi.org/10.5194/ESSD-15-4295-2023

Monteil, G., Broquet, G., Scholze, M., Lang, M., Karstens, U., Gerbig, C., Koch, F.-T., Smith, N. E., Thompson, R. L., Luijkx, I. T., White, E., Meesters, A., Ciais, P., Ganesan, A. L., Manning, A., Mischurow, M., Peters, W., Peylin, P., Tarniewicz, J., … Walton, E. M. (2020). The regional European atmospheric transport inversion comparison, EUROCOM: first results on European-wide terrestrial carbon fluxes for the period 2006–2015. *Atmos. Chem. Phys.*, *20*(20), 12063–12091. https://doi.org/10.5194/acp-20-12063-2020

Monteil, G., & Scholze, M. (2021). Regional CO2 inversions with LUMIA, the Lund University Modular Inversion Algorithm, v1.0. *Geosci. Model Dev.*, *14*(6), 3383–3406. https://doi.org/10.5194/gmd-14-3383-2021

Munassar, S., Monteil, G., Scholze, M., Karstens, U., Rödenbeck, C., Koch, F. T., Totsche, K. U., & Gerbig, C. (2023). Why do inverse models disagree? A case study with two European CO2 inversions. *Atmospheric Chemistry and Physics*, *23*(4), 2813–2828. https://doi.org/10.5194/ACP-23-2813-2023

Petrescu, A. M. R., McGrath, M. J., Andrew, R. M., Peylin, P., Peters, G. P., Ciais, P., Broquet, G., Tubiello, F. N., Gerbig, C., Pongratz, J., Janssens-Maenhout, G., Grassi, G., Nabuurs, G. J., Regnier, P., Lauerwald, R., Kuhnert, M., Balkovič, J., Schelhaas, M. J., van der Gon, H. A. C. D., … Dolman, A. J. (2021). The consolidated European synthesis of CO2emissions and removals for the European Union and United Kingdom: 1990-2018. *Earth System Science Data*, *13*(5), 2363–2406. https://doi.org/10.5194/ESSD-13-2363-2021

Thompson, R. L., Broquet, G., Gerbig, C., Koch, T., Lang, M., Monteil, G., Munassar, S., Nickless, A., Scholze, M., Ramonet, M., Karstens, U., van Schaik, E., Wu, Z., & Rödenbeck, C. (2020). Changes in net ecosystem exchange over Europe during the 2018 drought based on atmospheric observations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1810), 20190512. https://doi.org/10.1098/rstb.2019.0512

matrix formulation of the transport

The model estimate y^m for the observations y is given by

$$y^m = \mathcal{H}(x, E, y^{bg})$$

where:

- E represents the emissions of CO_2 and $C\Delta 14C$. For CO_2 those correspond to the F terms, defined in Section 2.1 of the paper, while for some of the $C\Delta 14C$ emission categories, the definitions are slightly more complicated, for instance, the $C\Delta 14C$ fossil fuel emissions are given by $E_{ff}^{C\Delta} = \Delta_{ff} F_{ff}$ (see Equation 4b of the paper for the definition of all the categories). The emissions are provided at a 0.5°, hourly resolution. Therefore E can be seen as an array of dimensions ($ncat_{mod}$, nt_{mod}, $nlat$, $nlon$). Note that there is no "tracer" dimension: some emission categories refer to CO_2, while some others refer to $C\Delta 14C$.
- y^{bg} represents the background concentrations (i.e. the influence of emissions further away in time and/or space). Note that "background" here should be understood from a modelling perspective (it is a form of boundary condition), and not in a data assimilation sense.
- x is the control vector, which is adjusted by the inversion. It contains a correction to the emissions E, in the form of offsets. x is provided at a lower resolution than E, and has size $n_{opt} = ncat_{opt} \times nt_{opt} \times np_{opt}$, with n_{opt} the number of optimised emission categories (2 or 3 depending on the simulations), nt_{opt} the number of weekly optimisation time steps, and np_{opt} the number of optimised "clusters" (or patches, regions, super-cells, ...), as shown in Figure 2b of the paper.
- $\mathcal{H}$ is the observation operator, which links the control vector x, containing the adjusted by the inversion (i.e. in the optimisation space), to the corresponding values in the observation space. It handles essentially three operations:
    1. project the coarse resolution emission offsets x onto the model grid
    2. compute the regional transport of these emissions, i.e. their influence on observed values
    3. combine it with the background concentrations y^{bg}

In our case, the atmospheric transport is linear, meaning that the previous equation can be re-written as

$$y^m = y^{bg} + \mathbf{K}(\mathbf{e} + \mathbf{P}\mathbf{x})$$

with:

- $\mathbf{P}$ the (n_{mod}, n_{opt}) matrix, projecting $\mathbf{x}$ onto the emission space.
- $\mathbf{e}$ the vectorized form of E (i.e. $\mathbf{e}$ is a $nmod = ncat_{mod} \times nt_{mod} \times nlat \times nlon$ vector)
- $\mathbf{K}$ is the transport matrix, projecting the emissions onto the observations space. $\mathbf{K}$ has shape ($nobs$, $nmod$)

Note that with this formulation, when $\mathbf{x}(:)$ is 0, (i.e. when $\mathbf{x} = \mathbf{x_b}$), the equation simplified to $y^m_{apri} = y^{bg} + \mathbf{K}\mathbf{e}$: The original, fine-resolution structure (0.5°, hourly) of the emissions is always

present and transported, even when the control vector is at a much lower resolution.

Classically, inversions seek to minimise a cost function $J(\mathbf{x})$ that balances the fit to the prior $\mathbf{x_b}$ with the fit to observations y:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x_b})^T\mathbf{B^{-1}}(\mathbf{x}-\mathbf{x_b}) + \frac{1}{2}(\mathbf{y_m}-\mathbf{y})^T\mathbf{R^{-1}}(\mathbf{y_m}-\mathbf{y})$$

We then define $\delta y = y - y^m_{apri}$. Since $y^m_{apri} = y^{bg} + \mathbf{Ke}$, we can replace $y^m - y$ in the equation above by simply $\mathbf{Kx} - \delta y$:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x_b})^T\mathbf{B^{-1}}(\mathbf{x}-\mathbf{x_b}) + \frac{1}{2}(\mathbf{KPx}-\delta\mathbf{y})^T\mathbf{R^{-1}}(\mathbf{KPx}-\delta\mathbf{y})$$

This means that our observation operator $\mathbf{H} = \mathbf{KP}$ is fully linear, and has an adjoint $\mathbf{H^T} = \mathbf{P^T K^T}$.

The transport operator $\mathbf{K}$ is computed using the Lagrangian transport model FLEXPART, and consists essentially of one FLEXPART footprint for each observation. The operator $\mathbf{P}$ is simply regridding the offsets $\mathbf{x}$ onto the model grid. This is done through a series of matrix operations:

$$\mathbf{e}^{apos} = \mathbf{e}^{apri} + \delta\mathbf{e}$$

with:

$$\delta\mathbf{e_c} = (\mathbf{T_H} \otimes \mathbf{T_T})\mathbf{x_c}$$

where $\mathbf{x_c}$ is the portion of the control vector $\mathbf{x}$ containing the offsets to the emissions in category c. $\mathbf{T_T}$ is a (nt_{mod}, nt_{opt}) matrix such that $\mathbf{T_t}(i,j)$ contains the fraction (between 0 and 1) of the optimisation time step j that falls within the model time step i (so, in our case, it typically contains either 0 (when the model interval i is outside the optimisation interval j), or 1 / 168 (i.e. 1 / 7 / 24), when the model interval i is within the optimization interval j (since the model is hourly and the optimisation is weekly). Likewise $\mathbf{T_H}$ is a (np_{mod}, np_{opt}) matrix (with $np_{mod} = nlat \times nlon$) such that $\mathbf{T_H}(i,j)$ contains the fraction of the area of the optimisation cluster j that falls within the model grid cell i. The values range between 0 and 1 since the optimisation "grid" is irregular (Figure 2b).

The formula above is not practical ($\mathbf{T_H} \otimes \mathbf{T_T}$ is very large!), but it can be reformulated as

$$\delta e_c = vec(\mathbf{T_T X_c T_H^T})$$

, where $\mathbf{X_c}$ is $\mathbf{x_c}$ reshaped as a (nt_{opt}, np_{opt}) matrix, and vec is the operator reshaping the result of that operation from a (nt_{mod}, np_{mod}) matrix to a $nt_{mod} \times np_{mod}$ vector.

The operation is conducted in a similar way for each category. Mathematically, this can be formulated

$$\delta\mathbf{e} = (\mathbf{T_c} \otimes \mathbf{T_H} \otimes \mathbf{T_T})\mathbf{x_c}$$

with $\mathbf{T_C}$ a $(ncat_{mod}, ncat_{opt})$ matrix such that $\mathbf{T_c}(i,j)$ is 1 if the category i is optimised, and 0 if it is not.

In practice, the calculations are performed one observation at a time, and one category at a time:

$$y^i_{apos} = y^{bg} + \sum_c \mathbf{K}^i\mathbf{e_c} + \sum_{c_{opt}} \mathbf{K^i T_t X_{c_{opt}} T_h}$$

, where c is the list of categories relevant for the tracer (CO_2 or $C\Delta 14C$) corresponding to the observation y^i, and c_{opt} is the list of optimised categories relevant for that tracer (i.e. *ff* and *bio* for CO_2, plus *biodis* for $C\Delta 14C$, in some simulations). Each row of $\mathbf{K}$ corresponds to one FLEXPART footprint, out of which only the non-zero components are stored (but, mathematically, these footprints are still defined over the entire length of the simulation). Furthermore, $\mathbf{K}$ is in fact identical for all categories (so its real shape is only ($nobs$, $nt_{mod} \times nlat \times nlon$)).