Response to Reviewers

Reviewer comments are in black.

Author responses are in blue.

Changes made to the manuscript are in red.

Reviewer #2 (RC2)

This study describes the generation of a dataset using the newest version of FLEDXPART(V11), along with the new version HAMSTER(V2) for postprocessing, enabling the attribution of moisture and sensible heat source. The authors introduced the configuration of FELXPART in domain-filling mode to obtain air parcels trajectories, and present analyses of global moisture sources for both continental and oceanic precipitation, as well as global sensible heat sources of diabatic temperature increments.

In addition to these climatological global analyses, two regional cases are also examined. The usability of the dataset is demonstrated for different research applications. Overall, the manuscript is well organized and clearly written. The dataset is accessible and can also be used with tools other than HAMSTER. The main area for improvement lies in enhancing the clarity of the post-processing (HAMSTERV2) and the structure of the dataset.

General comments

1. HAMSTERV2 is a tool that reads the FELXPART-ERA5 dataset and post-processes it to identify the moisture sources of precipitation or precipitable water via evaporation, as well as the sensible heat sources associated with potential temperature increments. This dataset can also be analysed using other tools for a broader range of research questions; therefore, a clear and detailed description of the dataset is essential.

To my understanding, the dataset is generated from forward simulations of FLEXPART driven by ERA5 as input. Approximately 20 million air parcels are released each year, each air parcel with a unique identifier. Outputs are provided every 3 hours, including the position of each air parcel (latitude, longitude, height) and other associated variables.

Although the primary focus of this study is not the configuration of FLEXPART itself, it is still necessary to clarify some key settings. For example, the simulations are conducted on a year-by-year basis, (i.e., including one additional month from both the preceding year and subsequent years), but how new air parcels are released.

For users applying Lagrangian approaches but not familiar with the detailed implementation of FLEXPART, these aspects are difficult to understand without further clarification.

Author response: We thank the reviewer for this thoughtful comment. We agree that, as the dataset can be used independently of HAMSTER and with other analysis frameworks, a clear and sufficiently detailed description of the FLEXPART configuration and dataset structure is essential, particularly for users less familiar with Lagrangian particle dispersion modelling. While some aspects of the dataset structure were previously described in the Data Availability section, we acknowledge that key elements of the FLEXPART setup—such as parcel initialization and simulation design—were not sufficiently detailed in the Methods section.

Changes: We have expanded Sect. 2.1 to provide a clearer and more structured description of the dataset and FLEXPART configuration. In particular, we have:

- Added an explicit description of the parcel initialization strategy, clarifying that all parcels are initialized at the start of each simulation period in a domain-filling configuration rather than being continuously released. (L145–146)

  "All parcels are initialized at the start of each simulation period and are uniformly distributed to represent the global atmospheric mass, rather than being continuously released from specific source regions."

- Clarified that each parcel is uniquely identified and tracked throughout the simulation (L146–147).

  "Each parcel is assigned a unique identifier and tracked throughout the simulation."

- Added a description of the stored output, specifying that parcel positions and associated thermodynamic properties are recorded at 3-hourly intervals (L154–156).

  "At regular output intervals (3-hourly), the position and associated (thermo)dynamic properties of each parcel (e.g., latitude, longitude, height, temperature, specific humidity, and density) are stored, forming the basis of the trajectory dataset."

- We also clarified the temporal limitation imposed by the simulation design (L140–142).

  "Due to this year-by-year simulation setup, parcel trajectories are temporally bounded by the simulation period (December of the previous year to January of the following year) and cannot be extended further across calendar years."

These additions aim to improve transparency and make the dataset more accessible to a broader community.

2. The second main component of the manuscript illustrates how the dataset can be used in global and regional studies of moisture contribution to precipitation to precipitation and sensible heat sources to diabatic temperature increments. Taking moisture source attribution as an example: for one air parcel, the specific humidity is used to diagnose moisture uptake or loss when passing through a grid cell, which is straightforward. However, an additional “bias-correction” step is applied, in which the specific humidity change is adjusted proportionally according to the evaporation flux at each time step. This adjusted increase is then interpreted as the moisture contribution from that grid cell.

In principle, the evolution of specific humidity in an air parcel reflects evapotranspiration and moisture exchange among air parcels. Although the bias-correction ensures that the sum of E2P equals precipitation, it may distort the spatial pattern of E2P, potentially leading to an overestimation of moisture contribution from nearby regions and consequently a higher local recycling ratio. This is consistent with Fig.A1a in Appendix A, which suggests that this Lagrangian framework tends to overestimate contributions over continents.

For sensible heat tracking, the situation is more complex, as the change in diabatic temperature is also influenced by precipitation processes, introducing additional uncertainty compared with the moisture source attribution.

A discussion of uncertainties could also be included in the manuscript.

Author response: We agree that the bias-correction step modifies the diagnosed moisture tendencies and can, in principle, influence the spatial distribution of inferred source contributions (Introduction L90-92).

As mentioned in Appendix A (Fig. A1), the Lagrangian attribution may exhibit a systematic tendency to overestimate continental contributions and underestimate remote ones. To address these issues, HAMSTER v2 applies an *en route* bias correction that constrains parcel humidity changes to surface fluxes. This removes the artificial penalisation of older moisture contributions present in earlier WaterSip-type methods, where negative changes in specific humidity, often arising from the entrainment of dry air into the boundary layer, were misinterpreted as precipitation losses and distant sources were consequently underestimated. In addition, extending the backtracking period to 30 days enables a larger fraction of precipitation to be traced to its origins and allows remote moisture contributions to be recovered more accurately.

We therefore interpret the bias-correction as improving the physical consistency of the diagnosed source–sink relationships rather than amplifying local recycling. As a consequence, the resulting continental recycling ratios are more consistent with independent Eulerian estimates and generally lower than those reported by other Lagrangian frameworks.

We further agree that uncertainties remain, particularly for sensible heat attribution, where multiple diabatic processes contribute to potential temperature changes. This limitation is already discussed in Appendix B. More generally, uncertainties associated with the attribution framework are addressed in the limitations section and Appendix B of the manuscript.

Changes: Added a more detailed description of this overestimation of local sources in Appendix B (L472–482).

"This residual bias reflects the tendency of Lagrangian attribution to overestimate local contributions and underestimate remote sources, consistent with previous findings in WaterSip-based approaches (Cloux et al., 2021). A key reason is that negative changes in specific humidity along trajectories are commonly interpreted as precipitation losses, although they may also arise from processes such as the entrainment of dry air into the boundary layer. In such cases, remote moisture contributions can be artificially discounted too strongly, leading to a potential overestimation of nearby sources and an underestimation of distant ones (Cloux et al., 2021; Li et al., 2024b). The applied en route bias-correction procedure mitigates this effect by constraining parcel humidity tendencies to surface fluxes, unlike HAMSTER v1 (Keune et al., 2022), where the bias correction was applied only after attribution, while the longer backtracking period allows a larger fraction of long-range moisture transport to be captured. Together, these changes lead to a more balanced representation of source–sink relationships, although the bias is not fully removed."

Specific comments

Line64-46: "The second approach, introduced by Stohl and James (2004), and used in models such as the FLEXible PARTicle dispersion model (FLEXPART) (Stohl and James, 2004; Bakels et al., 2024), follows air parcels while explicitly tracking their evolving specific humidity and other state variables." This description may require clarification as FLEXPART is a Lagrangian particle dispersion model used to derive air parcel trajectories and variables along trajectories, and may not directly correspond to either of the two moisture tracking frameworks discussed here. Actually, in the "Bakels et al., 2024" paper, this second approach is not mentioned.

Author response: We thank the reviewer for this important clarification. We agree that FLEXPART itself is a Lagrangian particle dispersion model used to generate air parcel trajectories, rather than a moisture tracking framework in the same sense as the two methodological classes discussed. Our original wording may have conflated the trajectory model with the diagnostic framework applied to it.

Changes: We have revised the sentence to clearly distinguish between the trajectory model and the moisture tracking methodology (L65–70). It now reads:

"The second approach, introduced by Stohl and James (2004), follows air parcels while explicitly tracking their evolving specific humidity and other state variables along trajectories. This avoids the well-mixed assumption, though changes in parcel humidity reflect only the net balance of evaporation and precipitation and thus cannot separate

the two when they occur simultaneously. In practice, this approach is often implemented using Lagrangian particle dispersion models such as the FLEXible PARTicle dispersion model (FLEXPART, Stohl and James, 2004; Bakels et al., 2024), which provide the underlying air parcel trajectories and associated variables.”

Line66-67: The evolution of specific humidity represents the net balance between evaporation and precipitation; therefore, their individual contributions cannot be fully disentangled when they occur simultaneously. More importantly, the change in specific humidity of an air parcel is not only related to evaporation and precipitation within the current grid cell, but may also be influenced by moisture flux/convergence from other grid cells. This point may need to be clarified.

Author response: We thank the reviewer for this helpful comment. We agree that, in an Eulerian framework, moisture convergence can increase the moisture content of a fixed grid cell or control volume, for example in frontal or convective regions where moist air is brought together by converging flow. However, this interpretation should be distinguished from the Lagrangian framework used here. In a parcel-following moisture budget, advective effects are accounted for through the motion of the parcel itself, rather than through an additional convergence term. Thus, convergence can lead to a local accumulation of moist air in a fixed region, but it does not by itself act as a source of specific humidity for each individual parcel.

For example, in a frontal convergence region, a parcel does not remain stationary and accumulate moisture locally; it follows the three-dimensional flow, including the associated ascent. Changes in parcel specific humidity are therefore interpreted as the net effect of moisture uptake and loss processes acting along its trajectory. We have clarified this point in the manuscript and now explicitly state that the method provides a net source–sink estimate, and therefore cannot fully disentangle simultaneous uptake and loss processes occurring within the same time interval.

Changes: /

Line 68-69: In the sentence: “Both types of Lagrangian approaches can provide high-resolution, flow-following diagnostics of moisture transport without reliance on a fixed grid.” It is unclear whether this refers to higher temporal resolution compared to Eulerian approaches, i.e., the use of small time steps to obtain a high-frequency, flow-following diagnostics. If so, please consider rephrasing this sentence for clarity.

Author response: We thank the reviewer for pointing out this ambiguity. We intended to emphasize that Lagrangian approaches follow air parcels along their trajectories and are therefore not constrained by a fixed spatial grid, rather than referring to temporal resolution.

Changes: We have rephrased the sentence for clarity (L70–71): *“Both types of Lagrangian approaches can provide high-resolution, flow-following diagnostics of moisture transport without reliance on a fixed spatial grid.”*

Line 86-87: “Bias-correction requires global simulations and was shown to substantially affect both the magnitude and spatial structure of inferred source–receptor relationships and reduce associated uncertainties (Keune and Miralles, 2019).” I don’t completely understand why bias-correction scheme in HAMSTER requires global simulations. This is also not consistent with line 151-152 “a user-defined region”.

Author response: We thank the reviewer for pointing out this inconsistency. We agree that the statement was misleading and not always correct.

Changes: We removed “requires global simulations and” from the sentence (L90).

Line 91: The statement “The availability of hourly ERA5 fields” indicates that the ERA5 reanalysis dataset is used. However, in lines 121-122, “For the present dataset, flex_extract is used to download 3-hourly ERA5 fields at 0.5° horizontal resolution and convert them into FLEXPART-ready input files.” appears that you download the original ERA5 reanalysis and preprocess it for FLEXPART, is it what you mean? If so, please consider rephrasing the 121-122 sentence.

Author response: We indeed use ERA5 reanalysis data, which are retrieved and preprocessed using flex_extract before being used as input for FLEXPART.

Changes: We clarified the wording in Sect. 2.1 (L133-135). The sentence now reads:

“For the present dataset, ERA5 reanalysis fields are retrieved using flex_extract to download and preprocess the data into FLEXPART-ready input files at 3-hourly temporal resolution and 0.5° horizontal resolution.”

Line 102: “a full description of the HAMSTER v2 post-processing tool as open-access resources.” Section 2.2 provides the first detailed presentation of HAMSTER v2 in the manuscript. However, the structure of this section is not entirely clear, for example: “Source regions are identified by analysing changes in specific humidity (q) and potential temperature (θ) along each trajectory. For moisture, HAMSTER integrates evaporation minus precipitation (E − P) along the parcel path, where positive (negative) values indicate moisture uptake (loss).” These two sentences are confusing; how do you relate (E-P) with specific humidity change? Also these are not whole steps, in lines 170-175, the “bias correction” step is introduced, maybe consider putting the steps together. Reorganizing the method description may improve readability and clarity for readers.

Author response: We agree that this phrasing is not totally clear.

Changes: We rephrased the whole paragraph (L174–181), it now reads

"HAMSTER traces all air parcels present over a user-defined study region backward in time to determine when and where they originally gained moisture and heat via evaporation and sensible heat flux, respectively. Source regions are identified by analysing changes in specific humidity q and potential temperature (θ) along each trajectory. For moisture, changes in q are interpreted as net moisture tendencies along the parcel path, where positive (negative) values indicate moisture uptake (loss). For sensible heat (H), analogous changes in potential temperature, calculated from temperature, specific humidity and density, are used to diagnose surface heat uptake (Keune et al., 2022). The algorithm is based on the WaterSip approach, introduced by Sodemann et al. (2008), which uses linear discounting so that earlier evaporative or sensible heat sources contribute less to the sink region when precipitation or heat losses occur along the trajectory. In backward mode, …"

Line 151–152: The use of "forward or backward" may need clarification. For moisture attribution, backward tracking is typically applied. If HAMSTER also supports forward tracking for diagnosing moisture and sensible heat sinks, this could be clarified.

Author response: We thank the reviewer for pointing this out. We agree that the original wording was misleading, as the description refers specifically to backward tracking, which is also the approach used throughout this study.

Changes:

- We have removed "forward or" from the sentence to clarify that the method description refers to backward tracking (L174).
- We have added a clarification in Appendix B noting that HAMSTER also supports forward tracking (L491–492):

  "Such forward analyses are also supported within the HAMSTER framework, enabling the diagnosis of downwind impacts of upwind surface fluxes. "

Line 159-160: H2T is derived from changes in potential temperature and is interpreted as the attribution of diabatic temperature tendencies along air-parcel trajectories to earlier surface sensible heat flux. However, potential temperature changes reflect the combined effects multiple diabatic processes, including both sensible and latent heating. It is therefore unclear how the contribution from sensible heat fluxes is separated from other heating processes, which may require further clarification.

Author response: We agree that changes in potential temperature reflect the combined effects of multiple diabatic processes, including both sensible and latent heating, and therefore errors are incurred when attributing these temperature changes solely to surface sensible heat fluxes based on trajectory information alone.

In HAMSTER, the diagnosed heat uptake should be interpreted as a proxy for diabatic heat uptake constrained by sensible heat fluxes. This is achieved through the bias correction procedure, which constrains the total heat uptake along trajectories using surface sensible heat flux estimates. In other words, the vertical sum of the positive changes in potential temperature across all parcels is constrained to match the surface sensible heat flux. This ensures that the total attributed amount is consistent with the prescribed surface fluxes, although the vertical and temporal location of the attributed heat uptake along the trajectories may remain uncertain and may be influenced by other diabatic processes, such as latent heat release. We are working on a framework to separate and attribute heat sources more precisely with HAMSTER but that is outside the scope and objective of this study.

This limitation, together with the absence of a bias correction for heat losses, is inherent to Lagrangian heat tracking approaches and is already discussed in Appendix B.

Changes: Stated it explicitly and referred to Appendix B in Sect. 2.2 (186–187).

"However, potential temperature is also influenced by latent heating and radiative cooling, which cannot be separated in the present framework and should be kept in mind when interpreting these results (see Appendix B)."

Line 175: "Consequently, when only precipitation bias correction was applied, the correction reduced to a uniform scaling of the E2P field and did not affect the relative contributions of the different source regions." It is unclear whether this step should be interpreted as a bias-correction procedure, or rather as a normalisation enforcing consistency with precipitation constraints. In the latter case, it would not alter the relative source attribution, but only rescale the E2P field. Clarification or rephrasing would be helpful.

Author response: We thank the reviewer for this clarification. We agree that, in the case where only precipitation bias correction is applied, the procedure effectively reduces to a uniform rescaling rather than altering relative source contributions.

Changes: We rephrased the sentence for clarity (L464–466):

"Consequently, when only precipitation bias correction was applied, the procedure effectively reduces to a uniform rescaling of the E2P field, enforcing consistency with precipitation constraints without altering the relative contributions of different source regions."

Line 277-278: “Similarly, H in Fig. 5a will equate the sum of the H2T from all sinks, with mild deviations from ERA5 data due to internal consistency (Appendix A).” However, Appendix A only presents the difference between the sum of evaporation and sum of E2P, rather than any heat-related quantities. It would be helpful to also provide a similar figure of heat. Given the complexity of heat processes, the discrepancy between H2T and the total sensible heat flux may be larger and, may not be fully explained by deviations from ERA5 data alone.

Author response: We agree that, in principle, a comparison analogous to Fig. A1 for heat would be valuable. However, the aggregation of H2T cannot be directly compared to surface sensible heat fluxes in the same way as E2P is compared to evaporation.

This is because H2T is diagnosed along backward trajectories and represents instantaneous values along the trajectory, rather than a cumulative quantity. As a result, contributions at successive trajectory time levels are not independent and cannot be summed in the same way as E2P, which represents accumulated moisture contributions and can be aggregated over time.

In addition, contributions at each time step reflect conditions along the parcel path and may extend across simulation boundaries (e.g. December contributions influencing trajectories in early January). Consequently, not all surface sensible heat fluxes within a given calendar year are fully represented within the corresponding tracking period, further complicating a direct comparison.

Changes: Added a sentence on this in Appendix A (L453–454).

“In contrast to moisture, H2T represents instantaneous contributions along trajectories, which prevents a direct comparison with annual surface sensible heat fluxes.”

Line431-433: “This pattern is consistent with previous studies reporting exaggerated local land contributions (Li et al., 2024b) and lower remote ones in other FLEXPART-based frameworks (Cloux et al., 2021).”

FLEXPART primarily represents the dynamical component of moisture transport rather than a complete framework for moisture source attribution. Therefore, it may be more appropriate to refer to “WaterSip-based” approaches here, particularly as HAMSTER is built upon the WaterSip approach? In addition, Cloux et al (2021) also clarify that FLEXPART itself is a Lagrangian particle dispersion model rather than a dedicated tool for moisture source analysis.

Author response: We thank the reviewer for this important clarification. We agree that FLEXPART itself is a Lagrangian particle dispersion model and not a moisture attribution framework. We also note that the reported biases cannot be attributed solely to either the transport model or the attribution method, but rather arise from their combined use. In particular, uncertainties in the dynamical component (e.g., representation of convergence and divergence) as well as assumptions in the attribution framework can both contribute to these patterns.

Changes: Rephrased sentence for clarity, it now reads (L441–444):

"This pattern is consistent with previous studies reporting exaggerated local land contributions (Li et al., 2024b) and lower remote ones in similar Lagrangian moisture attribution frameworks (Cloux et al., 2021), reflecting combined uncertainties in both the transport model and the attribution method."

Technical corrections

Line21: what's the meaning of "influence climate impacts", should it be "influence climate" or "have climate impacts on…"

Author response: We thank the reviewer for pointing out this ambiguity. We agree that the original phrasing was unclear.

Changes: Sentence has been revised to improve clarity. It now reads (L21):

"*Through these pathways, local surface conditions can have climate impacts far from their source."*

Line 62: The Dirmeyer and Brubaker method has also been used and tested in these studies:(Mu et al., 2026; Tuinenburg and Staal, 2020), so maybe consider mentioning these.

Author response: We thank the reviewer for this helpful suggestion. We agree that including these references strengthens the context around the application and evaluation of the Dirmeyer and Brubaker approach.

Changes: We have added the suggested references to the manuscript (L64–65).

"The first, following Dirmeyer and Brubaker (1999), diagnoses moisture sources from surface evaporation and precipitation fields … This approach has been widely applied and evaluated in subsequent studies (e.g., Tuinenburg and Staal, 2020; Mu et al., 2026)."

Line 151: "a user-defined the study region" -> "a user-defined study region".

Author response: We thank the reviewer for pointing out this typo.

Changes: The sentence has been corrected to *"a user-defined study region"* (L174).

Extra references added to manuscript:

Mu, Y., Evans, J. P., Taschetto, A. S., and Holgate, C.: Refining the Lagrangian approach for moisture source identification through sensitivity testing of assumptions using BTrIMS1.1, Geoscientific Model Development, 19, 1367–1385, https://doi.org/10.5194/gmd-19-1367-2026, 2026.

Tuinenburg, O. A. and Staal, A.: Tracking the global flows of atmospheric moisture and associated uncertainties, Hydrology and Earth System Sciences, 24, 2419–2435, https://doi.org/10.5194/hess-24-2419-2020, 2020.