Zhu et al. present a novel study that utilises the Nanjing University Carbon Assimilation System (NUCAS) v1.0 data assimilation framework and the process-based terrestrial ecosystem model Boreal Ecosystem Productivity Simulator (BEPS) is used as the adjoint model. The authors focus on simulating carbonyl sulfide (COS) fluxes at an ecosystem scale. Additionally, prior and posterior estimates of gross primary productivity (GPP), soil water content (SWC), sensible heat (H) and latent heat (LE) are presented. Assimilating COS flux measurements is a novel approach to better our understanding of COS processes, with regards to the carbon and water cycles, and the energy budget. This research is certainly within the scope of GMD and should be considered for publication following some minor corrections.

Having read the author's response to referee comments, Zhu et al. have clearly made a significant effort to improve their work. This is evidenced in the improved narrative of the paper and clarification on some of the methodology and data inputs. However, there is room for a bit more improvement in the readability, which my recommendations are focused on.

General Comments:

- The abstract would benefit from having at least one specific quantifiable metric that shows the improvements in COS flux or GPP made by NUCAS.

- There is use phrases like 'various constants' and other ambiguous phrases. I understand that elaborating in every instance could lead to excessive description of basic processes or repetition. But there is flexibility for some expansion on the use of these phrases. They often offer context and provide insight to readers who may not be familiar with this field, especially that of carbonyl sulfide which is rapidly becoming a field of its own within the world of carbon cycle science. An example on Lines 129-130: "The model parameters are the various constants that are not influenced by the model state."

- Sections 1 and 2 provide a thorough introduction to the topic and the methodology used in this work. However, the description of variables used in the modelling and assimilation is a bit light in places, leaving a lot of work for the reader to piece together. This harms to reproducibility of this work and makes it relatively inaccessible to a reader who is not familiar with matrix inversion, Bayesian statistics or data assimilation.
    - A sentence distinguishing biotic and abiotic soil processes would be helpful in Section 2.1.3.
    - Elaborate a bit on the parameters being referred to in Section 2.2.
    - Expand on the addition made in this iteration in Section 2.4.3, regarding coupling of COS with LE, H and SWC. It's still a bit light. At the very least include some additional references. Alternatively this could be elaborated on in the Discussion (Section 4), to which you could point the reader to from Section 2.4.3.
        - After giving this point some further thought, perhaps elaboration would be best suited to when the main scientific questions are presented in Section 1. I will leave this up to the discrepancy of the author(s).

- Section 3 is a thorough summary of the results. However, there are 2 key points that could be improved, firstly some more summary of the implication of what the parameters mean, i.e. what does it mean if the relative change in VJ_slope is large for example. Does it mean the assimilation has changed the posterior results significantly with regards to the prior and if so, what does that actually tell us? The summary at the end of 3.6 is a good example of where this

has been done well. A few more sentences like that would be good! Secondly and a much more minor point, there could be more specific reference to which figure and sub-figure is being referred to. It doesn't need to be excessive but try and help the reader from having to constantly check tables to cross-reference the PFT and soil texture between sites when this is referenced, for example. A good example of this occurring is at the start of Section 3.7, where you introduce the topic of this section and then just point the reader to Figure S4-S7. Provide a bit more guidance and refer to them as they are discussed in turn – I have included some specifics in the minor comments.

- Generally, I would say that the results for H, LE and SWC don't bring a huge amount to the paper. And large portions could be condensed or moved to supplementary material. But it certainly **does not** detract from the paper. However, it somewhat distracts the reader from what I would consider very good results regarding posterior COS fluxes and GPP. This is meant as an opinion and how to deal with this can be left to the discretion of the authors.

- There seems to be interchangeable use of "COS flux data", "COS flux measurements", "COS data" and "COS measurements". There is a big difference between a measurement of COS fluxes and ambient COS concentration for example. A good example is the first sentence of section 2.5.2: "After the ability of NUCAS to assimilate COS flux data was confirmed by twin experiments, we could then use the system was then utilised to conduct data assimilation experiments with real COS observations under single-site and multi-site conditions" – Is "real COS observations" referring to measurements of the flux? Please be diligent in distinguishing this throughout the paper. From what I can tell, the only use of ambient COS concentrations is to drive the estimation of prior COS fluxes.

- Some sentences are unclear. I have highlighted potential improvements in the readability of these in the Minor Comments. Please read them careful to ensure there has not been a misinterpretation.

Minor Comments:

- L22-24: improve readability. Example: "Data assimilation experiments were conducted to investigate the robustness of NUCAS, and to test the feasibility and applicability of assimilating carbonyl sulfide (COS) fluxes from seven surface sites, in order to better our understanding of stomatal conductance and photosynthesis."
- L26: I assume you mean COS fluxes: "assimilation of COS fluxes can".
- L28: Be consistent in referring to NUCAS as "the NUCAS" or simply "NUCAS". Also see note for Line 73.
- L28: "to the" -> "with"
- L36: "of earth system" -> "of the Earth system"
- L37: the biosphere 'significantly mitigating climate change' is a bit speculative in my opinion. It has certainly reduced the full potential of climatic changes since 1850. But the reasoning is partly through feedbacks, such as longer growing seasons in the Northern Hemisphere, as a result of climate change. I would just end the sentence at 1850 (and of course keep the reference).
- L38-39: "of terrestrial biosphere have changed" -> "of the terrestrial biosphere has changed".
- L41: "important tool to investigate" -> "important tool used to investigate".
- L44: "data" -> "datasets"

- L45: The reference the Scholze et al. (2017) feels a bit random. Be more specific about the 'various observations' you're referring to at the beginning of the sentence or just remove the reference. I would just start the sentence: "Observations such as sun-induced.."
- L47: Not sure this is a 'recent' finding any more. Remove.
- L52-54: Move these two sentences (starting 'Plants' and 'As' to the paragraph above), as they are still discussing the relationship between CO2 and COS. The sentence afterwards moves into discussing COS and GPP modelling.
- L58: What other 'key ecosystem variables' are you referring to? Are there any others that COS can estimate better than direct measurement? The reason it's useful to estimate GPP is because it is impossible to measure GPP directly at large spatial scales. Perhaps be specific about what you are referring to, the reader might be interested.
- L58: "However, with the lack of" -> "However, due to the lack of".
- L58-61: Move the references to Wohlfahrt et al. (2012) and Kooijmans et al. (2021) to the end of the sentence. Remove Bruhl et al. (2012), I don't understand the context of that reference here.
- L62: "behavior" -> "behaviour".
- L65: remove "various".
- L68-71: I think it's clear you're referring to data assimilation techniques. This sentence could be reduced and read better, example: "More specifically, the observed dynamics of ecosystems can be more accurately portrayed, additionally, our understanding of ecosystem processes can be deepened, with respect to their responses to climatic changes."
- L73: Include a note that NUCAS v1.0 will be referred to as NUCAS for the remainder of the paper: "NUCAS v1.0, hereafter referred to as NUCAS, is designed"
- L73-74: Improved punctuation. "to assimilate multiple observational data streams including COS flux data to improve the process based" -> "to assimilate multiple observational data streams, including COS flux data, to improve the process-based".
- L77-78: An example of non-line-breaking hyphen. This will be the only in-text mention I refer to. **Please see Technical Note on this.**
- L82: Clarify "COS fluxes"
- L85: Again, are you referring to COS observations or COS flux observations?
- L88: Combine and clean these sentences: "Materials and methods used in our study are described in Sect. 2, such as the BEPS model and NUCAS, are introduced".
- L89: What data do you mean? Just be a bit more specific: "along with the data used to drive BEPS and assimilated into NUCAS."
- L104: "if" -> "is" and "First" -> "first".
- L119: Remove BFGS acronym definition. It isn't used again.
- L128-132: Keep this paragraph with the previous one. The discussion is still relevant, i.e. regarding cost function.
- L140: Change "new" to "updated". I wouldn't refer to literature written in 1999 as new.
- L153-155: Be more specific about what is being referenced here: "The canopy-level COS plant uptake $Fcos,plant$ (pmol m^{-2} s^{-1}) was calculated by upscaling the resistance analog model of COS uptake, as presented by Berry et al. (2013), with the upscaling scheme recommended by Chen et al. (1999)."
- L161: No need to end the sentence. Also include another reference to Berry et al. (2013), as you are directly quoting the calculation: "where COS_a is the COS mole fraction in the bulk air and g_{sw} and g_{bw} are the stomatal conductance and leaf laminar boundary layer conductance to water vapor (H_2O), respectively (Berry et al., 2013)."

- L164-165: Try to avoid starting a sentence with a lower case, even if it is a relevant parameter. Perhaps: “The apparent conductance for COS uptake from the intercellular airspaces is denoted by g_{cos} and combines the mesophyll conductance and the biochemical reaction rate of COS and carbonic anhydrase (CA).”
- L166-169: Move the references to the end and lead the sentence into equation 4. Perhaps: “rate of Rubisco at 25℃ (Badger and Price, 1994; Evans et al., 1994), such that:
  Eq. (4)
  where α is a scaling parameter that is calibrated”
- L170: These values of α need a reference. If it’s from Stimler et al. (2012), then move the reference in the previous sentence to here.
- L170-171: “With reference to the COS modelling scheme: Simple” -> “According to the COS modelling scheme: Simple”
- L178: End this sentence with : rather than . As you are referring to Eq. 7 anyway.
- L180: Same as above. : rather than .
- L182: Now start sentence with where, instead of Where.
- L187: “Here” -> “In Eq. 10,” Also it is not clear what a is used for. Maybe add some context?
- L187: “maximum” -> “optimal”.
- L190: “reference of” -> “reference to”
- L191: This is the first use of SWC as an abbreviation, soil define it, i.e. soil water content (SWC).
- L195: use the variable names, “COS plant uptake” -> “$F_{cos,plant}$” and COS soil fluxes -> “$F_{cos,soil}$”. You can probably just remove this sentence to be honest. But up to the author.
- L198: “tuned previous model in development” -> “tuned in past model development”
- L200: Specifically refer to the research: “The prior uncertainty of parameters is set based on previous studies by Chen et al. (2022) and Ryu et al. (2018).”
- L222: “was” -> “were”
- L223-224: Sentence about LAI, meteorology and soil datasets need a reference. Or as they were mentioned in 2.1.2, perhaps direct the reader there.
- L226: “soil water content (SWC) at these sites collected at the sites were used” -> “SWC collected at the sites were used”.
- L231: “LAI product represents Lead area index at a” -> “LAI product quantifies leaf area index”.
- L231-233: spatial resolution is traditionally presented as 8 × 8 km for example. As it is a 2-D shape. Please clarify.
- L235: You may as well be specific in the products you are referring to. “The other two LAI products were used to investigate the effect of the LAI products” -> “The GLASS and MODIS LAI products were used to investigate the effect of different LAI products”
- L238-239: This last sentence is a bit wordy, but I understand what is to be communicated. Perhaps: “In addition, the 8-day temporal resolution of the LAI data was interpolated into daily values using the nearest neighbour method.”
- L241-245: These few sentences could be a bit neater and consistent. Perhaps: “Standard hourly meteorological data was inputted in BEPS, including air temperature at 2 m, shortwave radiation, precipitation, relative humidity and wind speed, taken from the FLUXNET database (for sites: AT-Neu, DK-Sor, ES-Lma, FI-Hyy and US-Ha1 see https://fluxnet.org), the AmeriFlux database (for sites: US-Ha1, US-Wrc, see https://ameriflux.lbl.gov) and the ERA5 dataset (for sites: AT-Neu, IT-Soy, US-Ha1 see https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=overview).”
- L247: Move the reference to just before the comma. Also remove “Particularly” at the start of the next sentence.

- L257-258: I think the eddy-covariance and gradient-based approach both require a reference.
- 258-259: “The COS soil measurements were collected using soil chamber, except at US-Ha1, where a sub-canopy flux-gradient approach was used to calculate the soil COS flux.” -> “The COS soil flux measurements were collected using soil chambers, except at US-Ha1, where a sub-canopy flux-gradient approach was used.”
- L260: “COS flux measurements”
- L262-264: The first sentence of this paragraph needs to be reworded. There is very little context, and the ordering is confusing. Also, what is the aerodynamic gradient method? IS this the same as the gradient-based approach? Perhaps: “US-Wrc utilises the gradient-based approach to measure COS ecosystem flux (reference), however available data is limited to only COS concentration measurements and lacking other parameters required, therefore this site risks introducing biases.” Or similar..
- L268-273: summary of Equation 11 could be much more concise. Perhaps: “This was done using the mean, $\bar{M}$, and standard deviation, σ_m, of the simulated COS flux to correct the COS flux observations (O):
  Eq. 11”
  where $\bar{O}$ and σ_O are mean and standard deviation of the observed COS flux series. F is the corrected observed COS flux and the COS simulations were calculated using the prior parameters for the time period corresponding to the COS flux observations.
- L273: No need to start new paragraph. Move up to previous.
- L273: “as an estimate”
- L276: “LE data are selected” -> “LE, data were selected”
- L277: This needs a reference.
- L283: Move this reference to the end of the sentence.
- L306-307: “With reference” -> “Regarding”
- L309: “After the ability of NUCAS to assimilate COS flux data was confirmed by twin experiments, we could then use the system” -> “After the ability of NUCAS to assimilate COS flux data was confirmed by twin experiments, the system was then utilised”
- L312: ~~of these~~
- L317-318: put comma (,) after experiment and simultaneously.
- L325: “in twin experiment” -> “in a twin experiment”.
- L326: was calculated “using Eq. 12”.
- L328: “where” -> “Where”.
- L352: I only see D_{final} in Table S5.
- L355: “nearly zero with the maximum value below” -> “nearly zero, where the maximum value was below”.
- L356: pseudo-observations
- L358: ~~results during the assimilation process~~.
- L364: ~~4.87%~~ I read this as 6.35% in Table 2?
- L365: “with the cost function reduction of 16.39% and 15.70%.” -> “~~with the cost function reduction~~ of 16.39% and 15.70% respectively.”
- L366: Include the percentage values for these 2 sites: “FI-Hyy (21.47%)” and “US-Wrc (27.71%)”.
- L367-368: I believe you mean July of 2015. This sentence needs to be reworded, perhaps: “In August 2014 and July 2015, the cost function reduction was between 40.59 % and 50.94 %, while in July of all other years, the cost function reduction was much lower, ranging from 4.87 % to 18.94 %.”

- L371-372: This is a bit of a throwaway sentence. If prior simulations were that good, we wouldn't need an inversion scheme right. I don't think sentence is necessary and almost devalues the posterior results.
- L377: I found this to be a particularly interesting finding. In that the way the COS fluxes are being calculated in the posterior are clearly missing one or more processes to exactly replicate measurements. Certainly something to investigate in future research and perhaps highlight in your conclusions.
- L382: You can remove the sentence starting "Similar to". I believe you raised this in the previous sentence.
- L384: 6.94 ~~pmol m−2 s−1~~ to 3.09 pmol m^{-2} s^{-1}
- L387: missed a t: nighttime.
- L390: "FI-Hyy and US-Wrc have different soil textures, with sandy loam and loam, respectively." -> "FI-Hyy and US-Wrc have different soil textures; sandy loam and loam respectively."
- L391: "took this difference into account and" -> "accounted for this difference appropriately and".
- L392: I calculate this to be 26.28% but perhaps I have mixed up numbers. Please check.
- L408: If the two-site assimilation method achieved similar results to the single-site, why do we need the two-site? Was is it more of proof-of-concept? A sentence summarising why it was useful would be helpful.
- L410: regarding 'as mentioned before', it looks like this material has been moved to the appendix. Please check and update accordingly.
- Section 3.4: It isn't really clear if a positive or negative change is a good thing. Especially as the majority of the summary refers to absolute differences. Could this be elaborated on?
- Line 419: 45.09% surely this value is hugely skewed by ES-Lma? What is the value excluding this site?
- L425-427: my interpretation of IT-Soy is that on paper the RMSE is ok and improved in the posterior due to improvements during the daytime. However, it's minimal change from the prior suggests it is not particularly sensitive to assimilation of COS flux data.
- L430: remove capital T.
- L432: include at the end of this sentence "(note the difference in x-axis scales)". By eye it is misleading initially.
- L435-438: These values are very different. Is it appropriate to be comparing variables like-for-like in this way? A bit more explanation of the implications of the results would be helpful.
- L442: Unless you have specifically excluded DK-Sor, I would remove this.
- L444-445: [at end of sentence] respectively.
- L447: Maybe this is clearer: "Our results also suggest that f_leaf tends to play a more important role in the COS assimilation at the forest sites (DK-Sor, FI-Hyy, US-Ha1 and US-Wrc) compared to the low-stature vegetation type sites (AT-Neu, ES-Lma and IT-Soy), with the mean absolute SIs about two times than that of the latter, with the exception of DK-Sor." Optional.
- L445: Does a lower R^2 value not suggest that the assimilation has worsened the result? Also do you mean Figure S3?
- Section 3.6: Lots of plots being referred to. Include 'see Figure 6c' etc.. where necessary. Help the reader.
- L465: I calculate the 3.81% to be 8.61%. Please check.
- L471: GPP? Not COS.
- L472: Drop line after 'underestimated.'
- L484: struggling to get where these 2 percentages have come from.

- L495: Please refer to figures S4-S7 as they are discussed. Rather than just listing them at the start of a section. Also as a note, if you are having to discuss and refer to figures in supplementary material, it's probably a sign that you are trying to present too much. As mentioned earlier, you could probably remove sections 3.7 and 3.8.
- L551: Requires more references, bottom-up or top-down. Kooijmans et al. (2021), Ma et al. (2021), Maignon et al. (2021) and Remaud et al. (2022). For example.
- L567-569: "COS plant uptake is governed by the hydrolysis reaction of COS (Wohlfahrt et al., 2012), catalysed by CA, though it can also be degraded by other photosynthetic enzymes, e.g., RuBisCo (Lorimer and Pierce, 1989), and the reaction is not dependent on light (Stimler et al., 2011; Whelan et al., 2018)." I think reads a bit better. Optional.
- L586: Proven.
- L593: CA, not carbonic anhydrase.
- L595: "capable to influence" -> "capable of influencing"
- L599: "sensitivity of V_{cmax25}" -> "sensitivity in V_{cmax25}"
- L608-611: It's not clear if you're saying your work also found this. Please clarify and amend accordingly.
- L623-624: "In comparison, the RMSEs of GPP simulations were reduced by an average of 25.37% within the assimilation of COS, while that of LE were reduced by 16.27 %." -> "In comparison, the RMSEs of GPP simulations were reduced by an average of 25.37 % as a result of assimilating COS, but reducing LE by only 16.27 %."
- L630: at the end of sentence: "via evapotranspiration".
- L635: behaviour
- L639: 'remarkable differences' is an odd phrase. Be specific. Also if you mean large, this is a different narrative to Section 3.
- L641: BEPS
- L679: move reference to end of sentence.
- Some of the discussion in Section 4.5 is a bit wordy (mainly last paragraph). Below are a few instances of trying to improve readability and flow.
    - Final paragraph, sentence 2: "As the nighttime COS plant uptake is driven by stomatal conductance (Kooijmans et al., 2021), g_n, nighttime COS fluxes can therefore be used to test the capability of BEPS to model g_n.
    - L699: space between 1 and mm.
    - Final paragraph sentence 5: "Similar findings by Resco De Dios et al. (2019), showed that the median g_n in the global dataset was 40 $\mathrm{mmol\ m^{-2}\ s^{-1}}$."
    - Final paragraph sentence 8: "As different enzymes have different physiological characteristics, Cho et al. (2023) proposed a new temperature function for the CA enzyme and showcased the considerable difference in temperature response of enzymatic activities of CA and RuBisCo, which provided valuable insights into the modelling and assimilation of COS."
    - L707: CA
    - L709: N = nitrogen?
    - L710: in -> by
    - Final paragraph final sentence: "Therefore, using the global microbial C biomass, soil N content and MAP datasets, the relationships between these variables, and the associated COS exchange processes, it is to be expected that a more accurate modelling of terrestrial ecosystem COS fluxes could be achieved, further increasing our understanding of the global COS budget and facilitate the assimilation of COS fluxes."

- L724-726: Perhaps: “Fourteen twin experiments, thirteen single-site experiments and one two-site experiment covering the period from 2012 to 2017, were conducted to investigate the capability of NUCAS to assimilate COS fluxes and optimize output parameters and variables. COS flux observations from a range of ecosystems were used, including four PFTs and three soil textures.”
- L729: COS fluxes
- L749: “throughout the plant function type” -> “for different PFTs”.
- L753-754: “However, the ‘equifinality’ can be avoided by imposing additional observational constraints (Beven, 2006).” Such as? i .e., ‘in this instance, we refer to the calculation and assimilation of multiple datasets, other than just COS fluxes.’.

Technical Notes:

- Be consistent with the following phrases single-site, two-site, AmeriFlux, process-based, etc.. in terms of the use of hyphen and capitalisation. Please check other potential sources of inconsistency.

- Use non-line-breaking hyphen where possible. This way be resolved in the editing by EGU, but for future reference to avoid a hyphenated phrase, or unit, breaking a line, use ctrl+shit+-.

- % symbols should be immediately adjacent to values, not with a space.