Reviewers' and Editors' comments:

We thank the Executive Editor and Associate Editor for processing our initial submission and giving us a chance to revise our manuscript. We also would like to thank the reviewers for reviewing our paper and providing their many useful comments and suggestions. We believe the revised version of the paper addresses their key concerns.

We now provide our detailed, point-by-point responses to the reviewers' comments. Their comments are in black font and our responses are in blue font. The modifications addressing the reviewers' comments in the revised version of the paper are tracked. We also made changes to this manuscript to improve the quality and readability.

Reviewer #2:

This manuscript discusses some interesting topics for geologic CO2 sequestration. Monitoring is important for long-term security and success of GCS projects, and optimization is a difficulty. Although the overall scope fits the broader scope of Applied Energy, the manuscript should be much improved to become a good quality article, especially for some obvious mistakes. Specific comments and suggestions are listed below:

Thank you very much for your insights and comments. We have modified the manuscript to reflect your revisions and improve its quality to make it suitable for publication in Applied Energy.

(1) Page6 Model Description: "Permeability distribution equal to 1 × 10 -1 m2 and 1 × 10 -13 m2, respectively". This is an obvious mistake of this very high permeability for caprock.

We apologize for this obvious oversight on our part. We have corrected the caprock permeability accordingly ($1 \times 10^{-19}\ m^2$).

(2) Page 7: The pressure at the model top is 0.2 MP, which is equal to about 10 m below the surface. This means that your reservoir is about 100 m below the surface, which is unreasonable for a GCS reservoir (at least 800 m or below). And is the "Aquifer" a USDW or a saline aquifer? The whole model setup is unrealistic.

We apologize for this oversight on our part as well. The pressure at the top of the aquifer zone is 9.81 MPa, corresponding to the 1000-meter depth of the reservoir below the surface (as describe in Page 6) and the assumed pressure gradient of $9.81 \times 10^{-3}\ MPa/m$. We meant to say that there is an additional overburden of 0.2MPa at the top of the reservoir zone in addition to the normal pressure gradient. This has been clarified in the manuscript.

Regarding the aquifer zone, we have clarified in the manuscript that this is a potential USDW aquifer, while the reservoir itself is a saline aquifer. We understand that this synthetic model might not be representative of real-world field cases, however we believe that the proposed method shows significant promise and can be adapted with some modifications to a real-world field case.

(3) Page 8: What is the reason for the assignment of data assimilation error tolerance? Is 0.002 MPa equivalent to 0.05 of CO2 saturation and 0.002 C for temperature?

Thank you for your inquiry. The data assimilation error tolerance, τ, is selected based on engineering judgement, as described in the last paragraph of Page 8. These are site-specific and subject to engineering judgement. In our case, we selected the τ for pressure, CO_2 saturation and temperature empirically based on the variance of the observed measurements. We have clarified this in the manuscript.

(4) Figure 10: The histograms of "Prior", "posterior_R1" and "posterior_R100" cannot be compared in the same figure - the width of each bar is not equal. They should be set to the same interval of each bar.

Thank you for this insight. In Figure 10, we plot the prior uncertainty in cumulative CO_2 leakage (blue) along with the posterior uncertainty for 2 randomly-selected realizations (R1 in orange and R100 in red). It is correct that the width of the bins do not represent the same scale for the prior and posterior distributions; however, we carefully select an **equal number of bins** in both prior and posteriors, to show the reduction in uncertainty (U_R) after the data assimilation. We believe that this is a good visualization of U_R because if we were to have equal-width bins for prior and posterior, then the posterior would only have a very small number of bins since the spread is significantly reduced. We have clarified Figure 10 in the manuscript to explain our reasoning and how to interpret this figure.

(5) This manuscript is very similar to Chen et al., (2018) https://www.sciencedirect.com/science/article/pii/S0306261918307372. What is the improvement and what is the novelty of this manuscript?

Thank you for pointing this out, we appreciate your observation. This is indeed a continuation of Chen et al. (2018), where we developed a framework for ensemble data assimilation based on a deep learning proxy for CO_2 monitoring. The key difference is that in Chen et al. (2018), the two scenarios have arbitrary legacy well locations and arbitrary monitoring well locations. In this work, we do not assume fixed monitoring well locations, but instead wrap the workflow on an optimization loop to determine the best placement for monitoring wells as well as the optimal monitoring measurement to assimilate for uncertainty reduction. For instance, Figure 3 shows the complete workflow, and while Chen et al. (2018) implement Steps 1-5 for monitoring data assimilation using a proxy, we extend this to estimate the optimal monitoring well location and measurement type by looping Steps 2-5 and obtaining the optimal values for an optimal uncertainty reduction through Step 6. We have modified several portions of the manuscript to clarify and emphasize these differences and the significant improvement and novelty provided in this paper. Thank you again for this insight.

Reviewer #3:

The manuscript presents an approach to geologic CO2 sequestration using a filtering-based data assimilation coupled with Artificial Neural Networks (ANNs) for developing reduced-order models. These models aim for computational efficiency grounded in full-physics numerical simulations of CO2 injection in saline aquifers. The authors highlight the method's application in two scenarios focusing on CO2 leakage through abandoned wellbores and illustrate the substantial uncertainty reduction in cumulative CO2 leakage at the geologic CO2 sequestration site.

Thank you for your comment. We appreciate your time and effort in understanding and revising this manuscript.

Note: we appreciate the fact that you supplied citations in your comments to existing literature that use similar methods as ours. However, when we received the reviewer comments, these have been omitted and we are unable to tract the citation that you are referring to. We have worked to comprehensively examine the literature and evaluate relevant citations for our paper.

The paper is well-structured and tackles a topic of significant interest. However, several areas require attention and improvement:

1. The manuscript needs to articulate its novel contributions more distinctly. Many aspects of the proposed workflow, such as ANN, reduced-order modeling, and MCMC methods, are well-established in the field. It is important to delineate clearly how this work diverges from or improves upon existing methods cited in the literature, including works by (see the work of : <<<Specific reference recommendations are being omitted. The authors should make a deep analysis of the specialized literature and use it to improve the paper>>>. The authors need to discuss in more clarity their new contribution to the literature.

We appreciate your comment and expertise in this area. We have modified several portions of the manuscript, including the Introduction where we discuss relevant literature, to reflect the contribution and novelty of this manuscript. Thank you.

In Page 2, paragraphs 4 (“Efforts have been made to select…”), 5 (“Similar efforts have been made…”), and 6 (“Besides optimal well placement and monitoring measurement…”) we describe relevant literature in the areas of geologic CO_2 storage (GCS) monitoring design. Paragraph 4 describes efforts regarding optimal monitoring measurements to quantify the uncertainty in leakage risks. On the other hand, Paragraph 5 describes efforts made for optimal monitoring well placement in GCS. Paragraph 6 describes efforts in uncertainty quantification for GCS.

We have improved Paragraphs 1 and 2 in Page 7 to describe our contributions to these areas.

2. The manuscript's density of references, particularly in the introduction, is excessive and detracts from its readability and focus. While comprehensive citation is valuable, the current approach borders on overwhelming, with more than 130 references, many of which are not directly relevant to the core contributions of the paper. Streamlining the references to focus on the most relevant works will improve the manuscript's clarity and impact. I suggest reducing the reference by a factor of 2 at least.

Thank you for this suggestion. We agree that the density of references is significant and have reduced the number of references to only the most significant and relevant ones. We value the clarity and readability of our paper and appreciate your suggestion.

3. The authors need to discuss the applicability of their proposed approach for real-field cases. The selected case described in section 2.4 represents a synthetic scenario with given reservoir dimensions and subsurface fluid and rock properties, as well as, a predefined leakage scenario. The presented model is only applicable to the selected model.

Thank you for this insight. We have included in the Discussion the limitations of our method as it applied to synthetic geologic models with a given dimension and rock-fluid properties. We discuss the need for re-training or modifying the architecture in order to be applicable for real-world field cases.

4. A comparative discussion against analytical-based methods such as those presented would provide a more comprehensive evaluation of the proposed method's performance and limitations. Such analysis can help in understanding the conditions under which the proposed method might offer significant advantages or face challenges. Some of the proposed analytical-based methods based on pressure transient analysis are :
<<<Specific reference recommendations are being omitted. The authors should make a deep analysis of the specialized literature and use it to improve the paper>>>

We appreciate your comment and suggestion regarding a comparison against analytical-based methods. As we noted above, we are unable to read your suggested citation.

We agree that a comparison against analytical-based methods would be useful for the practitioner to decide when and how to use different techniques for GCS monitoring design. However, it is out of the scope of this manuscript to compare different techniques. Nonetheless, we have included comments in the Discussion and Conclusion regarding potential advantage and disadvantages of our method and potential comparisons against analytical-based methods.

5. The methodological section requires more depth, particularly in explaining the neural network architecture, training process, and how it integrates with the filtering-based data assimilation approach. Specific details on the data used for training, validation, and testing, along with performance metrics, will add substantial credibility and allow for a thorough understanding of the method's robustness.

Thank you for your comment, we appreciate credibility and thorough understanding of our proposed ANN ROM method. The entire Section 2.2 describes the experimental design, forward simulations and collection of training data, and model training and validation, as well as a more detailed description of the ANN ROM architecture, training, and performance. Figures 1 and 2 support Section 2.2 as well. – We have included additional details regarding the ANN ROM in Section 2.2 for additional clarity and depth.

Section 2.3 and Figure 3 describe the entire optimal monitoring design workflow, including a detailed description of how the ANN ROM integrates with the filtering-based data assimilation approach.