GEOPRIOR-MAMBA: STRUCTURED PROCESS PRIORS WITH MAMBA FOR FINE-RESOLUTION XCO_2 RECONSTRUCTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Reconstructing fine-resolution column-averaged dry-air CO_2 (XCO_2) fields from sparse satellite observations requires models to infer spatial structure that is only weakly constrained by direct measurements. Existing learning-based methods typically treat environmental covariates as ordinary numerical inputs and must therefore learn heterogeneous source–sink relationships largely from sparse supervision. We introduce GeoPrior-Mamba, a multi-directional Mamba framework augmented with offline language-model-induced structured process priors. Rather than using a language model to predict XCO_2, we use it before training to organize relative process knowledge for biospheric uptake, ecosystem respiration, and anthropogenic emissions into deterministic prior tables. These priors are spatially instantiated using geographic, ecological, emission-related, and seasonal information and are adaptively injected into the reconstruction backbone through a lightweight knowledge adapter. Using OCO-2 observations from 2018–2020, GeoPrior-Mamba achieves an RMSE of 0.81 ppm and an R^2 of 0.93 on held-out observations, reducing RMSE by 48.2% relative to CAMS background interpolation and by 3.1% relative to Trans-XCO_2 under the same evaluation protocol. Ablation experiments show a measurable contribution from the knowledge-prior branch and substantially faster convergence than the knowledge-free Mamba backbone. Independent TCCON evaluation further supports the consistency of the reconstructed fields with ground-based column CO_2 measurements. These results suggest that language models can provide a practical mechanism for constructing structured process priors when globally consistent process-response representations are difficult to obtain directly, while remaining outside the numerical prediction loop.

1 INTRODUCTION

Satellite measurements have greatly expanded the spatial coverage of atmospheric carbon dioxide observations, but they remain sparse and irregular in space and time. For column-averaged dry-air CO_2 (XCO_2), valid retrievals from missions such as OCO-2 are limited by cloud and aerosol contamination, surface reflectance, illumination geometry, orbital sampling, and retrieval quality control. Consequently, satellite observations alone do not provide spatially continuous fine-resolution XCO_2 fields. (Eldering et al., 2017; Pandey et al., 2024; Yang et al., 2025)

Reconstructing XCO_2 from incomplete observations is challenging because large parts of the target field are only weakly constrained by direct measurements. Existing statistical, assimilation-based, and learning-based methods address this problem in different ways. Data-driven methods are especially flexible because they can combine atmospheric background fields, meteorological variables, vegetation indices, emissions, geographic information, and temporal signals. However, these predictors are usually treated as ordinary numerical channels. The model must therefore infer relationships associated with biospheric uptake, ecosystem respiration, anthropogenic emissions, and their geographic and seasonal dependence largely from sparse satellite supervision. Representative reconstruction studies are discussed in Section 2.1, with a broader comparison provided in Appendix D. (Gao et al., 2023; He et al., 2022; Li et al., 2023; Wang et al., 2023; Zhang et al., 2023b)

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Many of these relationships are already partly understood at the process level, but globally consistent and spatially complete representations of such process-response knowledge are not readily available from existing observational or model products. The challenge is therefore to organize dispersed scientific knowledge into a deterministic, computable form that is spatially aligned with gridded data and compatible with end-to-end learning. Large language models provide one possible mechanism for this organization. We therefore ask a different question: can language-model knowledge be converted offline into structured numerical priors that guide a geoscientific reconstruction model? (Zhang et al., 2024; Zscheischler et al., 2017)

We propose GeoPrior-Mamba, which separates knowledge organization from numerical prediction. An offline language model is used to construct fixed relative-activity priors for major XCO_2 source–sink processes under different geographic and ecological conditions. These priors are then instantiated as pixel-wise fields and adaptively injected into a multi-directional Mamba reconstruction backbone. The language model is never used to predict XCO_2 directly and is not called during model training or inference.

Our contributions are threefold:

1. We introduce a framework for converting offline language-model knowledge into deterministic, spatially instantiated process priors for sparse geoscientific reconstruction.
2. We develop GeoPrior-Mamba, which adaptively integrates these priors with a shared multi-directional Mamba backbone for daily global XCO_2 reconstruction.
3. We evaluate the framework using OCO-2 observations from 2018–2020, showing improved held-out reconstruction relative to CAMS background interpolation and Trans-XCO_2, measurable gains over a knowledge-free Mamba backbone, faster convergence, and broad consistency with independent TCCON observations.

2 Related Work

2.1 XCO_2 Reconstruction from Sparse Observations

Existing XCO_2 reconstruction methods include statistical interpolation, data-assimilation products, and learning-based approaches that combine satellite retrievals with auxiliary environmental predictors. Recent learning-based methods improve spatial completeness and resolution, but they generally represent meteorological, vegetation, emission, and geographic information as ordinary numerical inputs. GeoPrior-Mamba is complementary to this line of work: its main distinction is not a new source of supervision, but an explicit mechanism for injecting structured source–sink process context into a sparse reconstruction model. Representative examples include statistical and ensemble reconstruction, self-supervised fusion, high-resolution deep learning, and recent global daily products (Gao et al., 2023; Guan et al., 2024; He et al., 2022; Li et al., 2022; Wang et al., 2023; 2025; Zhang et al., 2023b). A more extensive literature comparison is provided in Appendix D.

2.2 Knowledge-Guided Scientific Machine Learning

Knowledge-guided and physics-informed learning incorporate scientific structure through governing equations, simulation outputs, expert rules, constraints, or structured priors. Our setting differs from approaches that directly enforce physical equations because fine-resolution XCO_2 reconstruction does not provide a complete closed-form process model at the target scale. Instead, we use structured relative-activity priors as soft contextual guidance and allow the neural network to learn when and where they are useful. (Raissi et al., 2019; Zscheischler et al., 2017)

2.3 Language Models as Scientific Knowledge Sources

Language models are increasingly used to organize scientific information, generate hypotheses, retrieve domain knowledge, or support scientific workflows. GeoPrior-Mamba uses a more constrained role: the language model is queried only offline to construct fixed numerical priors, and no language-model component is present during training or inference. This design reduces computational overhead and avoids making numerical predictions depend directly on stochastic language-model outputs. (Zhang et al., 2024)

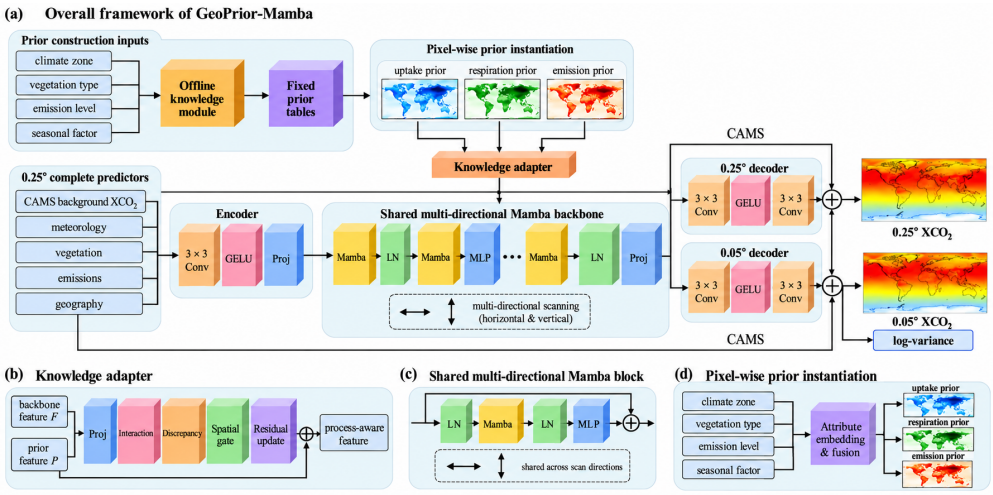

Figure 1: Architecture of the proposed GeoPrior-Mamba reconstruction model.

3 Problem Setup and Data

We formulate XCO_2 reconstruction as supervised learning with sparse observational constraints. The model receives complete gridded predictors on a daily 0.25° input grid and produces a spatially continuous XCO_2 field on a 0.05° target grid. OCO-2 observations mapped to the 0.05° grid provide supervision only where valid retrievals are available; unobserved grid cells do not contribute to the supervised reconstruction loss.

The predictors include a CAMS XCO_2 background field, ERA5 meteorological variables, MODIS vegetation indices, ODIAC fossil-fuel emissions, and static geographic attributes. CAMS provides the large-scale atmospheric background, while meteorological, vegetation, emission, and geographic variables provide information related to transport, surface conditions, biospheric activity, anthropogenic forcing, and spatial heterogeneity. Climate zone, vegetation type, and emission level are additionally used to instantiate the structured process priors described in Section 4. Independent TCCON observations are used only for external evaluation and are never used for model training.

Table S1 summarizes the datasets and their roles in the reconstruction framework. Detailed product citations, preprocessing procedures, and native resolutions are provided in Appendix A.

4 Method

4.1 Framework Overview

As shown in Figure 1, GeoPrior-Mamba contains three components. First, an offline language-model-based knowledge module organizes process knowledge for biospheric uptake, ecosystem respiration, and anthropogenic emissions into fixed relative-activity priors. Second, these priors are mapped to each grid cell and date using geographic, ecological, emission-related, and seasonal attributes. Third, a knowledge adapter injects the resulting process fields into a multi-directional Mamba backbone that reconstructs spatially continuous XCO_2.

This design explicitly separates knowledge elicitation, prior instantiation, and numerical reconstruction. The language model is used only before training, while the reconstruction network remains fully deterministic during training and inference.

4.2 Offline Construction of Structured Process Priors

For biospheric processes, the language model is queried for combinations of climate zone and vegetation type to obtain candidate relative-activity scores for vegetation uptake and ecosystem respiration. These values are normalized process indicators rather than estimates of physical carbon

flux. Anthropogenic forcing is represented by a separate prior indexed by emission level, while vegetation-dependent seasonal factors modulate the biospheric priors across reconstruction dates.

To reduce sensitivity to individual language-model outputs, multiple independent responses are generated for each attribute combination. Candidate values are screened using predefined format checks and geoscientific sanity constraints, including bounded non-negative activity values and seasonally reasonable vegetation behavior. Valid responses are aggregated using the median. The resulting prior tables are fixed before model training and remain unchanged thereafter.

The language model is therefore used to organize process knowledge rather than to simulate carbon fluxes or provide direct XCO_2 targets. Its role is not to replace observational or model-derived flux products, but to provide a scalable way to construct a globally consistent process-prior representation from dispersed process-response knowledge.

4.3 Pixel-Wise Prior Instantiation

The fixed prior tables are converted into spatially aligned process fields for each reconstruction date. Climate zone and vegetation type determine the baseline biospheric context of each pixel, emission level determines the anthropogenic prior, and vegetation-dependent seasonal factors modulate the biospheric fields through time. The resulting maps represent normalized relative activity associated with biospheric uptake, ecosystem respiration, and anthropogenic emissions.

These fields do not prescribe the reconstructed XCO_2 value. Instead, they provide structured context that the neural network may use selectively through the knowledge adapter.

4.4 Multi-Directional Mamba Backbone and Knowledge Adapter

The complete 0.25° predictor fields are first encoded into latent spatial features. To capture long-range spatial dependencies, the encoded features are processed along four spatial scan directions, including left-to-right, right-to-left, top-to-bottom, and bottom-to-top. A shared Mamba module is reused across these directions to limit parameter redundancy and reduce sensitivity to a single scan order. The directional representations are then combined to form the backbone feature representation. (Gu & Dao, 2023)

The structured process priors are projected into the latent feature space and spatially aligned with the backbone features. The knowledge adapter integrates the two representations through feature interaction and discrepancy modeling, and generates a process-aware feature update. A spatial gate controls the strength of prior injection at each grid cell, allowing the model to adaptively exploit process information while suppressing it where the data-driven representation provides stronger evidence. The resulting update is added to the backbone representation through a residual connection, preserving the original features while introducing structured process context.

The fused representation is subsequently decoded at two spatial scales. A coarse decoder produces the 0.25° reconstruction, while a fine-resolution decoder generates the final 0.05° XCO_2 field. Both outputs are formulated as residual corrections to the CAMS background field. In addition, the fine-resolution branch predicts a spatially varying log-variance field used in the uncertainty-aware training objective described in Section 4.5.

4.5 Training Objective

GeoPrior-Mamba is trained using sparse OCO-2 supervision together with spatial regularization. Reconstruction losses are evaluated only at grid cells with valid OCO-2 observations on the 0.25° and 0.05° grids. The model predicts residual corrections relative to the CAMS background field. In addition to the masked reconstruction loss, we use a Gaussian negative log-likelihood term for the predicted log variance and spatial regularization terms that constrain the smoothness and local gradients of the residual field.

The overall training objective is

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{rec}} + \lambda_{\mathrm{nll}}\mathcal{L}_{\mathrm{nll}} + \lambda_{\mathrm{lap}}\mathcal{L}_{\mathrm{lap}} + \lambda_{\mathrm{grad}}\mathcal{L}_{\mathrm{grad}}. \tag{1}$$

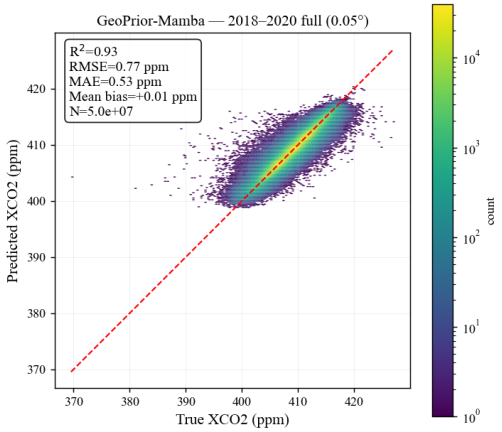

Figure 2: Comparison between 0.05° reconstructed XCO_2 and gridded OCO-2 reference XCO_2 during 2018–2020.

Here, $\mathcal{L}_{\mathrm{rec}}$ is the masked reconstruction loss, $\mathcal{L}_{\mathrm{nll}}$ denotes the uncertainty-aware Gaussian negative log-likelihood, and $\mathcal{L}_{\mathrm{lap}}$ and $\mathcal{L}_{\mathrm{grad}}$ regularize the spatial smoothness and local gradients of the predicted residual, respectively. Detailed loss definitions are provided in the appendix.

5 Experimental Setup

5.1 Data Split and Evaluation

We use OCO-2 observations from 2018–2020 and apply a date-based split so that all valid observations from held-out reconstruction dates are excluded from training. This evaluates the ability of the model to reconstruct XCO_2 on unseen dates rather than merely interpolate among observations from the same day.

Performance is evaluated only at valid held-out OCO-2 grid cells using MSE, RMSE, MAE, bias, and R^2. Independent TCCON observations are collocated with the reconstructed fields in space and time for external validation. (Wunch et al., 2011)

5.2 Baselines and Ablations

We compare GeoPrior-Mamba with CAMS background interpolation and Trans-XCO_2 under the same split and observation mask. We also evaluate a knowledge-free Mamba variant to isolate the contribution of the structured prior branch. Additional ablations remove the 0.25° branch and the NLL term. (Meng et al., 2026)

6 Results

6.1 Main Reconstruction Performance

As shown in Figure 2, the reconstructed XCO_2 agrees well with the gridded OCO-2 reference observations over the full 2018–2020 evaluation period. GeoPrior-Mamba achieves an RMSE of 0.77 ppm, an MAE of 0.53 ppm, an MSE of 0.60 ppm^2, a near-zero bias, and an R^2 of 0.93. The close correspondence between reconstructed and observed XCO_2 indicates that the model captures most of the observed variability while maintaining low systematic error.

Table 1 summarizes the main reconstruction performance on the held-out test set. GeoPrior-Mamba achieves the best overall performance among the compared methods, with an R2of 0.925, an MAE of 0.558 ppm, an MSE of 0.658 ppm^2, and a near-zero bias of +0.002 ppm. The corresponding RMSE

Table 1: Main reconstruction performance and ablation results on the held-out test set.

Model	Held-out test set			
	$R^2 \uparrow$	MAE(ppm) ↓	MSE(ppm) ↓	Bias(ppm) ↓
Other Models				
CAMS interpolation (Baseline)	0.721	1.244	2.450	-0.307
Trans-XCO_2	0.921	0.579	0.699	+0.018
GeoPrior Mamba (ours)	**0.925**	**0.558**	**0.658**	**+0.002**
Ablation Experiments				
Without 0.25° branch	0.924	0.559	0.664	+0.002
Without knowledge branch	0.923	0.562	0.671	-0.019
Without NLL loss	0.924	0.562	0.666	+0.002

Table 2: Main reconstruction performance and ablation results during 2018–2020.

Model	2018-2020			
	$R^2 \uparrow$	MAE(ppm) ↓	MSE(ppm) ↓	Bias(ppm) ↓
Other Models				
CAMS interpolation (Baseline)	0.724	1.232	2.404	-0.320
Trans-XCO_2	0.927	0.552	0.632	-0.017
GeoPrior Mamba (ours)	**0.931**	**0.533**	**0.598**	**+0.007**
Ablation Experiments				
Without 0.25° branch	0.931	0.534	0.600	+0.008
Without knowledge branch	0.930	0.540	0.609	-0.016
Without NLL loss	0.930	0.537	0.605	+0.011

is 0.811 ppm. Relative to CAMS background interpolation, GeoPrior-Mamba reduces RMSE from 1.565 to 0.811 ppm, corresponding to a reduction of approximately 48.2%, while increasing R^2 from 0.721 to 0.925.

GeoPrior-Mamba also improves over Trans-XCO_2 under the same evaluation protocol, reducing MSE from 0.699 to 0.658 ppm^2 and RMSE from 0.836 to 0.811 ppm. Although the improvement over Trans-XCO_2 is smaller than that over the CAMS background, it is obtained relative to a substantially stronger learning-based baseline.

Table 2 further reports the corresponding performance over the 2018–2020 period. The results show a consistent pattern, with GeoPrior-Mamba achieving an R^2 of 0.931 and an MSE of 0.598 ppm^2, compared with 0.927 and 0.632 ppm^2 for Trans-XCO_2, respectively.

6.2 Independent Validation with TCCON

Table 3 summarizes the site-level validation against independent TCCON observations. The reconstructed XCO_2 fields show broad consistency with ground-based column CO_2 measurements. R^2 exceeds 0.90 at most evaluated sites, and RMSE is below 1 ppm at many stations, indicating that the reconstruction captures the temporal variability of XCO_2 across a wide range of locations. (Wunch et al., 2011)

Larger errors are observed at several sites, including Eureka and Xianghe. These deviations may reflect local atmospheric heterogeneity, representativeness differences between point-based TCCON measurements and 0.05° gridded reconstructions, and limited collocation samples at some locations. Overall, the TCCON comparison provides an independent evaluation beyond the OCO-2 observations used for model assessment.

6.3 Contribution of the Structured Prior Module

As shown in Tables 1 and 2, the ablation results exhibit a consistent pattern across the held-out test set and the 2018–2020 evaluation period. Removing the knowledge-prior branch produces the largest degradation among the tested components. On the held-out test set, MSE increases from 0.658 to 0.671 and R^2 decreases from 0.925 to 0.923, while the bias shifts from +0.002 to -0.019 ppm. A

Table 3: Site-level validation against TCCON observations during 2018–2020.

Site	n	R^2	MAE (ppm)	RMSE (ppm)	Bias (ppm)
Lauder 120HR (NZ)	102	0.806	0.269	0.356	+0.096
Lauder 125HR (NZ)	482	0.941	0.391	0.496	−0.077
Edwards (US)	785	0.952	0.457	0.581	−0.017
Burgos (PH)	493	0.955	0.527	0.661	−0.288
Wollongong (AU)	579	0.917	0.533	0.727	−0.308
Izana (ES)	346	0.942	0.643	0.766	+0.419
Lamont (US)	726	0.915	0.605	0.784	+0.052
Sodankylä (FI)	292	0.973	0.603	0.792	+0.397
Orleans (FR)	322	0.946	0.614	0.797	+0.054
JPL (US)	106	0.690	0.663	0.837	−0.070
Rikubetsu (JP)	210	0.958	0.692	0.874	+0.344
East Trout Lake (CA)	489	0.958	0.652	0.889	+0.072
Nicosia (CY)	398	0.940	0.732	0.889	+0.647
Paris (FR)	166	0.940	0.660	0.893	+0.092
Park Falls (US)	550	0.947	0.680	0.896	−0.068
Karlsruhe (DE)	274	0.928	0.721	0.933	+0.088
Pasadena/Caltech (US)	688	0.911	0.748	0.958	−0.450
Garmisch (DE)	255	0.938	0.790	0.961	+0.508
Saga (JP)	448	0.889	0.745	0.978	+0.323
Bialystok (PL)	122	0.933	0.778	0.986	−0.374
Bremen (DE)	93	0.942	0.818	1.072	+0.125
Tsukuba (JP)	281	0.868	0.831	1.086	−0.099
Ny-Ålesund (NO)	77	0.920	0.980	1.199	+0.908
Réunion (FR)	218	0.959	1.101	1.243	−0.956
Darwin (AU)	562	0.938	1.118	1.249	−1.103
Hefei (CN)	106	0.816	0.988	1.345	−0.161
Xianghe (CN)	463	0.845	1.348	1.753	−0.918
Eureka (CA)	55	0.901	2.064	2.226	+2.064

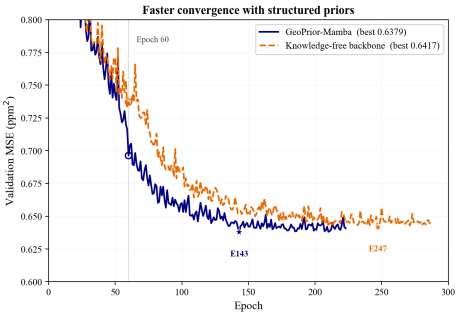

Figure 3: Validation MSE curves of GeoPrior-Mamba and the knowledge-free Mamba backbone. GeoPrior-Mamba reaches the best validation level of the knowledge-free backbone substantially earlier.

similar pattern is observed over 2018–2020, where removing the knowledge-prior branch increases MSE from 0.598 to 0.609 and reduces R^2 from 0.931 to 0.930.

Removing the 0.25° branch results in a smaller degradation, while excluding the NLL term also slightly increases reconstruction error in both evaluation settings. These results indicate that the Mamba backbone provides a strong reconstruction foundation, while the structured prior module contributes an additional and consistent improvement in accuracy and bias correction.

6.4 Convergence Behavior

Figure 3 compares the validation trajectories of GeoPrior-Mamba and the knowledge-free Mamba backbone. GeoPrior-Mamba reaches a validation MSE of 0.6379 at epoch 143, which is already lower than the best validation MSE achieved by the knowledge-free backbone (0.6417). The knowledge-free model requires 247 epochs to reach a comparable validation level, corresponding to 104 fewer optimization epochs, or approximately 42% fewer epochs for GeoPrior-Mamba.

The advantage is particularly evident during the intermediate stage of training. Around epoch 60, the prior-enhanced model reaches a validation MSE of 0.6963, whereas the knowledge-free backbone remains at 0.7328. Although the gap gradually narrows as training proceeds, GeoPrior-Mamba still converges to a lower validation error. This behavior suggests that the structured priors provide useful optimization guidance under sparse supervision.

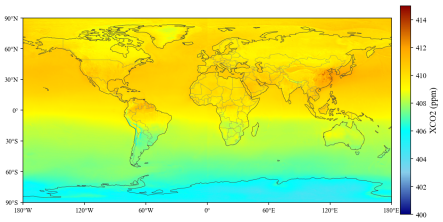

Figure 4: Spatial distribution of the multi-year mean XCO_2 during 2018–2020.

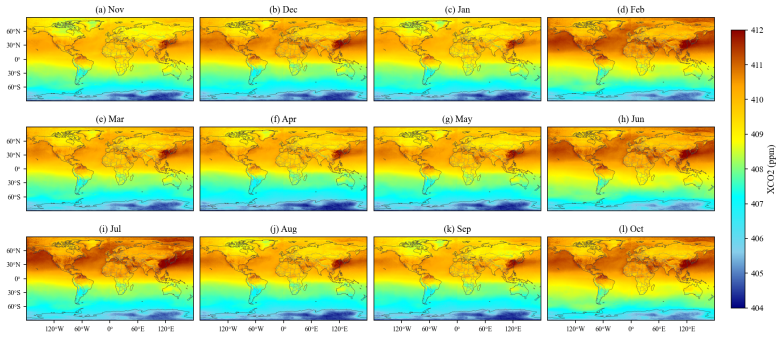

Figure 5: Global distributions of monthly mean XCO_2. Panels (a–l) correspond to November through October, arranged from left to right and top to bottom.

6.5 Spatial Characteristics of the Reconstruction

As shown in Figure 4, the multi-year mean reconstruction for 2018–2020 captures the dominant large-scale structure of atmospheric XCO_2, including higher concentrations in the Northern Hemisphere and lower concentrations toward the Southern Hemisphere. Regional spatial heterogeneity is also evident over major continental regions.

Beyond the large-scale background pattern, the reconstructed fields retain finer regional contrasts associated with heterogeneous surface and atmospheric conditions. These spatial variations are particularly evident over densely populated and strongly vegetated continental regions, where anthropogenic emissions and biospheric activity contribute to more pronounced local gradients in XCO_2. (Hakkarainen et al., 2016; Lei et al., 2021)

Figure 5 shows the multi-year monthly mean XCO_2 reconstructed by GeoPrior-Mamba over 2018–2020. A clear seasonal cycle is evident, particularly over the Northern Hemisphere. XCO_2 generally increases from boreal autumn into winter and remains relatively high from late winter through spring. Elevated concentrations are most apparent during approximately February–May, especially over the northern mid- and high-latitude continents. XCO_2 subsequently decreases during the boreal growing season, with lower concentrations appearing over northern continental regions in July–August, before increasing again from September toward winter. Despite these seasonal variations, a persistent interhemispheric gradient is maintained throughout the year, with generally higher XCO_2 in the Northern Hemisphere and lower values over the Southern Hemisphere and Southern Ocean.

The seasonal evolution is broadly consistent with the annual cycle of terrestrial carbon exchange in the Northern Hemisphere. Enhanced vegetation activity during the growing season is associated with stronger atmospheric CO_2 uptake and contributes to the summertime reduction in XCO_2, whereas weaker photosynthetic uptake outside the growing season is accompanied by a gradual recovery of atmospheric CO_2 concentrations. The seasonal contrast is particularly evident over Eurasia and North America, where the reconstructed fields show pronounced summertime drawdown relative to the preceding spring months. Regional enhancements also remain visible over densely populated continental regions, superimposed on the large-scale seasonal and interhemispheric patterns. (Graven et al., 2013; Zeng et al., 2014)

7 Discussion and Limitations

GeoPrior-Mamba differs from standard learning-based XCO_2 reconstruction primarily in how process knowledge is represented. The language model is not used as a predictor and does not generate XCO_2 fields. Instead, it is used offline to organize broad geoscientific knowledge into fixed relative-activity priors that are spatially instantiated and then interpreted adaptively by the reconstruction network. This separation preserves deterministic numerical inference while providing an explicit pathway for source–sink process knowledge to enter the model.

The current experiments indicate that the prior branch provides a measurable but moderate improvement in final reconstruction error and a clearer improvement in convergence behavior. These results support describing the priors as structured optimization guidance, rather than claiming that they dominate predictive accuracy. The stronger comparison is therefore between the knowledge-free Mamba backbone and GeoPrior-Mamba, while the comparison with Trans-XCO_2 reflects the combined effect of backbone and prior design. Importantly, the present experiments evaluate the downstream utility of these LLM-elicited priors rather than establishing the empirical equivalence of individual prior values to directly measured process quantities. This suggests that language models may serve as a practical knowledge interface for constructing usable global process priors when such process-response information is difficult to assemble consistently.

Several limitations remain. First, the model is primarily data-driven and does not explicitly solve atmospheric transport equations or enforce CO_2 mass conservation. The reconstructed fields should therefore not be interpreted as dynamically closed simulations. Second, the language-model-induced priors are simplified representations of relative process activity rather than physical carbon fluxes. Their values depend on the prompt design, aggregation procedure, and selected process categories. Finally, evaluation over longer periods and additional observational regimes is needed to assess generalization beyond 2018–2020.

8 Conclusion

We introduced GeoPrior-Mamba, a framework that converts offline language-model knowledge into deterministic, spatially instantiated process priors for sparse XCO_2 reconstruction. The priors encode biospheric uptake, ecosystem respiration, and anthropogenic emission context and are adaptively integrated into a multi-directional Mamba backbone. On OCO-2 reconstruction from 2018–2020, GeoPrior-Mamba improves over CAMS and Trans-XCO_2 under the same evaluation protocol and provides a measurable gain over a knowledge-free Mamba backbone. The prior-enhanced model also reaches comparable validation performance in substantially fewer optimization epochs, while independent TCCON evaluation supports the broad consistency of the reconstructed fields. These results suggest that language-model knowledge can be useful for numerical geoscientific learning when it is converted into fixed, constrained process representations rather than used directly in the prediction loop. Future work should incorporate transport- or mass-conservation-related constraints and evaluate generalization across longer periods and other sparse geoscientific reconstruction tasks.

AI USE STATEMENT

Generative AI tools were used in two roles. First, as part of the proposed methodology, a language model was used offline to organize geoscientific process knowledge into structured numerical priors. The language model was not used during model training or inference, and candidate prior values were screened, aggregated, and fixed before training. Second, generative AI tools were used to assist with language editing, LaTeX and formatting support, and preparation or refinement of the schematic architecture figure. They were not used to generate or alter the experimental data or reported numerical results. All AI-assisted text, figures, citations, and scientific claims were reviewed and verified by the authors.

REPRODUCIBILITY STATEMENT

The manuscript describes the data split, evaluation protocol, model architecture, and ablation design. The appendix provides detailed dataset and preprocessing information, complete loss definitions, the evaluation protocol, and an extended comparison with previous XCO_2 reconstruction studies. The structured-prior construction procedure, including repeated elicitation, screening, aggregation, and fixed-table instantiation, is described in Section 4.2.

REFERENCES

F. P. Antezana Lopez, G. Zhou, G. Jing, K. Zhang, L. Chen, L. Chen, and Y. Tan. Global daily column average CO2 at $0.1^\circ \times 0.1^\circ$ spatial resolution integrating OCO-3, GOSAT, CAMS with EOF and deep learning, 2025. Sci. Data, vol. 12, Art. no. 268, doi: 10.1038/s41597-024-04135-w.

S. Bhattacharjee and J. Chen. Prediction of satellite-based column CO2 concentration by combining emission inventory and LULC information, 2020. IEEE Trans. Geosci. Remote Sens., vol. 58, no. 12, pp. 8285–8300, doi: 10.1109/TGRS.2020.2985047.

B. Byrne, D. F. Baker, S. Basu, M. Bertolacci, K. W. Bowman, D. Carroll, et al. National CO2 budgets (2015–2020) inferred from atmospheric CO2 observations in support of the global stocktake, 2023. Earth Syst. Sci. Data, vol. 15, no. 2, pp. 963–1004, doi: 10.5194/essd-15-963-2023.

J. Chen, R. Hu, L. Chen, Z. Liao, L. Che, and T. Li. Multi-sensor integrated mapping of global XCO2 from 2015 to 2021 with a local random forest model, 2024. ISPRS J. Photogramm. Remote Sens., vol. 208, pp. 107–120, doi: 10.1016/j.isprsjprs.2024.01.009.

W. Cheng, L. Dan, X. Deng, X. Huang, et al. Global monthly gridded atmospheric carbon dioxide concentrations under the historical and future scenarios, 2022. Sci. Data, vol. 9, no. 1, Art. no. 83, doi: 10.1038/s41597-022-01196-7.

A. Eldering, P. O. Wennberg, D. Crisp, et al. The Orbiting Carbon Observatory-2 early science investigations of regional carbon dioxide fluxes, 2017. Science, vol. 358, no. 6360, Art. no. eaam5745, doi: 10.1126/science.aam5745.

P. Friedlingstein, M. O'Sullivan, M. W. Jones, et al. Global carbon budget 2022, 2022. Earth Syst. Sci. Data, vol. 14, no. 11, pp. 4811–4900, doi: 10.5194/essd-14-4811-2022.

Z. Gao, Y. Jiang, J. He, et al. Spatiotemporal variation analysis of global XCO2 concentration during 2010–2020 based on DINEOF-BME framework and wavelet function, 2023. Sci. Total Environ., vol. 892, Art. no. 164750, doi: 10.1016/j.scitotenv.2023.164750.

H. D. Graven, R. F. Keeling, S. C. Piper, et al. Enhanced seasonal exchange of CO2 by northern ecosystems since 1960, 2013. Science, vol. 341, no. 6150, pp. 1085–1089, doi: 10.1126/science.1239207.

A. Gu and T. Dao. Mamba: Linear-time sequence modeling with selective state spaces, 2023. arXiv:2312.00752.

X. Guan, Z. Sun, D. Chu, et al. Long-term (2000–2020) global 0.05° continuous atmospheric carbon dioxide mapping combining OCO-2 observations and model simulations, 2024. Sci. Total Environ., vol. 957, Art. no. 177051, doi: 10.1016/j.scitotenv.2024.177051.

J. Hakkarainen, I. Ialongo, and J. Tamminen. Direct space-based observations of anthropogenic CO2 emission areas from OCO-2, 2016. Geophys. Res. Lett., vol. 43, no. 21, pp. 11400–11406, doi: 10.1002/2016GL070885.

D. M. Hammerling, A. M. Michalak, and S. R. Kawa. Mapping of CO2 at high spatiotemporal resolution using satellite observations: Global distributions from OCO-2, 2012. J. Geophys. Res. Atmos., vol. 117, no. D6, doi: 10.1029/2011JD017015.

C. He, M. Ji, T. Li, et al. Deriving full-coverage and fine-scale XCO2 across China based on OCO-2 satellite retrievals and CarbonTracker output, 2022. Geophys. Res. Lett., vol. 49, no. 12, Art. no. e2022GL098435, doi: 10.1029/2022GL098435.

Q. He, Y. Liu, Z. Li, et al. Spatiotemporal dynamics of atmospheric CO2 across China revealed by long-term, high-resolution satellite-derived data, 2026. EGUsphere, pp. 1–23, doi: 10.5194/egusphere-2025-5647.

H. Hersbach, B. Bell, P. Berrisford, et al. The ERA5 global reanalysis, 2020. Q. J. R. Meteorol. Soc., vol. 146, no. 730, pp. 1999–2049, doi: 10.1002/qj.3803.

Y. Hua, X. Zhao, W. Sun, et al. Satellite-based reconstruction of atmospheric CO2 concentration over China using a hybrid CNN and spatiotemporal kriging model, 2024. Remote Sens., vol. 16, no. 13, Art. no. 2433, doi: 10.3390/rs16132433.

X. Huang, H. Yang, Q. Lv, et al. A full-coverage daily XCO2 dataset in China from 2015 to 2020 based on DSC-DF-LGB, 2024. Earth Syst. Sci. Data Discuss., pp. 1–22, doi: 10.5194/essd-2024-371.

A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. Overview of the radiometric and biophysical performance of the MODIS vegetation indices, 2002. Remote Sens. Environ., vol. 83, no. 1–2, pp. 195–213, doi: 10.1016/S0034-4257(02)00096-2.

S. Hwang, H. Choi, Y. Kang, et al. Reconstructing long-term (2003–2019) global high-resolution XCO2: Bridging observational gaps with machine learning, 2026. GISci. Remote Sens., vol. 63, no. 1, Art. no. 2627042, doi: 10.1080/15481603.2026.2627042.

A. Inness, M. Ades, A. Agustí-Panareda, et al. The CAMS reanalysis of atmospheric composition, 2019. Atmos. Chem. Phys., vol. 19, pp. 3515–3556, doi: 10.5194/acp-19-3515-2019.

C. Jin, Y. Xue, X. Jiang, et al. A long-term global XCO2 dataset: Ensemble of satellite products, 2022. Atmos. Res., vol. 279, Art. no. 106385, doi: 10.1016/j.atmosres.2022.106385.

R. Lei, S. Feng, A. Danjou, et al. Fossil fuel CO2 emissions over metropolitan areas from space: A multi-model analysis of OCO-2 data over Lahore, Pakistan, 2021. Remote Sens. Environ., vol. 264, Art. no. 112625, doi: 10.1016/j.rse.2021.112625.

J. Li, K. Jia, X. Wei, et al. High-spatiotemporal resolution mapping of spatiotemporally continuous atmospheric CO2 concentrations over the global continent, 2022. Int. J. Appl. Earth Obs. Geoinf., vol. 108, Art. no. 102743, doi: 10.1016/j.jag.2022.102743.

J. Li, Z. Zhang, T. Li, et al. Global daily seamless XCO2 mapping (2016–2020): Spatio-temporal trends and variations during wildfire events, 2026. Int. J. Appl. Earth Obs. Geoinf., vol. 146, Art. no. 105092, doi: 10.1016/j.jag.2026.105092.

T. Li, J. Wu, and T. Wang. Generating daily high-resolution and full-coverage XCO2 across China from 2015 to 2020 based on OCO-2 and CAMS data, 2023. Sci. Total Environ., vol. 893, Art. no. 164921, doi: 10.1016/j.scitotenv.2023.164921.

J. Liang, J. Cao, G. Sun, et al. SwinIR: Image restoration using swin transformer, 2021. in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), pp. 1833–1844.

W. Liu, R. Li, J. Cao, et al. Mapping high-resolution XCO2 concentrations in China from 2015 to 2020 based on spatiotemporal ensemble learning model, 2024. Ecol. Inform., vol. 83, Art. no. 102806, doi: 10.1016/j.ecoinf.2024.102806.

Z. Meng, J. Zheng, Y. Zhu, and H. Fu. Global daily XCO2 reconstruction at 0.05° resolution for 2021–2024 using satellite and reanalysis data, 2026. IEEE Trans. Geosci. Remote Sens., doi: 10.1109/TGRS.2026.3716990.

T. Oda and S. Maksyutov. A very high-resolution (1 km × 1 km) global fossil fuel CO2 emission inventory derived using a point source database and satellite observations of nighttime lights, 2011. Atmos. Chem. Phys., vol. 11, pp. 543–556, doi: 10.5194/acp-11-543-2011.

G. Pan, Y. Xu, and J. Ma. The potential of CO2 satellite monitoring for climate governance: A review, 2021. J. Environ. Manage., vol. 277, Art. no. 111423, doi: 10.1016/j.jenvman.2020.111423.

S. Pandey, J. B. Miller, S. Basu, et al. Toward low-latency estimation of atmospheric CO2 growth rates using satellite observations: Evaluating sampling errors of satellite and in situ observing approaches, 2024. AGU Adv., vol. 5, no. 4, Art. no. e2023AV001145, doi: 10.1029/2023AV001145.

Y. Qu, X. Shi, Y. Fan, et al. Reconstructing two-decadal global daily high-resolution XCO2 records using a hybrid Transformer–BiLSTM model, 2026. Earth Syst. Sci. Data Discuss., pp. 1–35, doi: 10.5194/essd-2026-50.

M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, 2019. J. Comput. Phys., vol. 378, pp. 686–707, doi: 10.1016/j.jcp.2018.10.045.

M. Reuter, M. Buchwitz, O. Schneising, et al. Ensemble-based satellite-derived carbon dioxide and methane column-averaged dry-air mole fraction data sets (2003–2018) for carbon and climate applications, 2020. Atmos. Meas. Tech., vol. 13, no. 2, pp. 789–819, doi: 10.5194/amt-13-789-2020.

M. Sheng, L. Lei, Z. C. Zeng, et al. Global land 1° mapping dataset of XCO2 from satellite observations of GOSAT and OCO-2 from 2009 to 2020, 2023. Big Earth Data, vol. 7, no. 1, pp. 170–190, doi: 10.1080/20964471.2022.2033149.

Y. Tang, J. Hu, D. Tang, et al. Spatio-temporal distribution of XCO2 concentrations in northeast China based on Downscaling-XGBoost model, 2024. in Proc. Int. Conf. Remote Sens., Mapping, Geograph. Inf. Syst. (RSMG), vol. 13402. Bellingham, WA, USA: SPIE, pp. 367–372, doi: 10.1117/12.3048860.

W. Tian, L. Zhang, Y. Yu, et al. Using multi-source data and time series features to construct a global terrestrial CO2 coverage by deep learning, 2024. IEEE Trans. Geosci. Remote Sens., doi: 10.1109/TGRS.2024.3462589.

J. Wang. Global daily 1 km gapless XCO2 (2003–2023) derived from multi-satellite observations and a spatiotemporal deep learning framework, 2026. Environ. Impact Assess. Rev., vol. 117, Art. no. 108146, doi: 10.1016/j.eiar.2025.108146.

Y. Wang, Q. Yuan, T. Li, et al. Seamless mapping of long-term (2010–2020) daily global XCO2 and XCH4 from the Greenhouse Gases Observing Satellite (GOSAT), Orbiting Carbon Observatory 2 (OCO-2), and CAMS global greenhouse gas reanalysis (CAMS-EGG4) with a spatiotemporally self-supervised fusion method, 2023. Earth Syst. Sci. Data, vol. 15, no. 8, pp. 3597–3622, doi: 10.5194/essd-15-3597-2023.

Z. Wang, C. Zhang, K. Shi, Y. Shangguan, B. Hu, X. Chen, D. Wei, S. Chen, P. M. Atkinson, and Q. Zhang. A full-coverage satellite-based global atmospheric CO2 dataset at 0.05° resolution from 2015 to 2021 for exploring global carbon dynamics, 2025. Earth Syst. Sci. Data, vol. 17, pp. 5355–5375, doi: 10.5194/essd-17-5355-2025.

C. Wu, Y. Ju, S. Yang, et al. Reconstructing annual XCO2 at a 1 km × 1 km spatial resolution across China from 2012 to 2019 based on a spatial CatBoost method, 2023. Environ. Res., vol. 236, Art. no. 116866, doi: 10.1016/j.envres.2023.116866.

D. Wunch, G. C. Toon, J.-F. L. Blavier, et al. The Total Carbon Column Observing Network, 2011. Philos. Trans. R. Soc. A, vol. 369, no. 1943, pp. 2087–2112, doi: 10.1098/rsta.2010.0240.

H. Yang, T. Li, J. Wu, et al. Inter-comparison and evaluation of global satellite XCO2 products, 2025. Geo-Spat. Inf. Sci., vol. 28, no. 1, pp. 131–144, doi: 10.1080/10095020.2023.2252017.

N. Zeng, F. Zhao, G. J. Collatz, et al. Agricultural Green Revolution as a driver of increasing atmospheric CO2 seasonal amplitude, 2014. Nature, vol. 515, no. 7527, pp. 394–397, doi: 10.1038/nature13893.

L. Zhang, F. Jiang, W. He, et al. A robust estimate of continental-scale terrestrial carbon sinks using GOSAT XCO2 retrievals, 2023a. Geophys. Res. Lett., vol. 50, no. 6, Art. no. e2023GL102815, doi: 10.1029/2023GL102815.

L. Zhang, T. Li, J. Wu, et al. Global estimates of gap-free and fine-scale CO2 concentrations during 2014–2020 from satellite and reanalysis data, 2023b. Environ. Int., vol. 178, Art. no. 108057, doi: 10.1016/j.envint.2023.108057.

M. Zhang and G. Liu. Mapping contiguous XCO2 by machine learning and analyzing the spatio-temporal variation in China from 2003 to 2019, 2023. Sci. Total Environ., vol. 858, Art. no. 159588, doi: 10.1016/j.scitotenv.2022.159588.

Y. Zhang, X. Chen, B. Jin, S. Wang, S. Ji, W. Wang, and J. Han. A comprehensive survey of scientific large language models and their applications in scientific discovery, 2024. in Proc. 2024 Conf. Empirical Methods in Natural Language Processing (EMNLP), pp. 8783–8817, doi: 10.18653/v1/2024.emnlp-main.498.

J. Zscheischler, M. D. Mahecha, V. Avitabile, et al. Reviews and syntheses: An empirical spatiotemporal description of the global surface–atmosphere carbon fluxes: Opportunities and data limitations, 2017. Biogeosciences, vol. 14, no. 15, pp. 3685–3703, doi: 10.5194/bg-14-3685-2017.

A Additional Data and Experimental Details

This appendix provides the dataset descriptions, preprocessing details, loss definitions, evaluation protocol, and an extended comparison with representative XCO_2 reconstruction studies. The study period is 2018–2020. Complete model predictors are prepared on a daily 0.25° grid, while OCO-2 supervision and the final reconstructed XCO_2 fields are represented at 0.05° resolution. Supervised reconstruction losses are evaluated only where valid OCO-2 observations are available.

A.1 Datasets and preprocessing

Table S1: Summary of the datasets used in this study.

Data	Variable	Temporal resolution	Spatial resolution	Source
OCO-2	XCO_2 retrievals	Irregular soundings; daily gridding	2.25 km × 1.29 km footprint; 0.05° grid	NASA GES DISC
CAMS background	Large-scale CO_2/XCO_2 background	Daily	Global field; resampled to 0.25°	Copernicus Atmosphere Monitoring Service
ERA5	Meteorological variables	Daily	0.25°	ECMWF/Copernicus CDS
MODIS	NDVI, EVI	16-day	1 km; resampled to 0.25°	NASA LP DAAC/MODIS
ODIAC	Fossil-fuel CO_2 emissions	Monthly	1 km; resampled to 0.25°	ODIAC/NIES
Topography	Elevation, slope, aspect	Static	Native DEM; resampled to 0.25°	Global digital elevation model
TCCON	Ground-based XCO_2	Site dependent; collocated in time	Station observations	TCCON Data Archive

OCO-2 supervision. Quality-controlled OCO-2 Level-2 XCO_2 retrievals from 2018–2020 are used as the primary supervisory observations (Eldering et al., 2017). Sounding-level retrievals are mapped to the daily 0.05° grid. When multiple valid soundings fall within the same grid cell on the same day, they are aggregated to obtain a representative grid-cell value. Grid cells without valid retrievals are treated as missing targets and do not contribute to the supervised reconstruction loss.

Background and environmental predictors. The CAMS global CO_2 field provides the large-scale atmospheric background used by the residual reconstruction framework (Inness et al., 2019). ERA5 variables include 2-m air temperature, 2-m dewpoint temperature, 10-m zonal and meridional winds, surface pressure, and precipitation, which characterize near-surface thermodynamic conditions and transport-related variability (Hersbach et al., 2020). MODIS NDVI and EVI represent vegetation density, seasonal greenness, and productivity-related variability (Huete et al., 2002). Monthly ODIAC

fossil-fuel CO_2 emissions characterize anthropogenic forcing, particularly over urban and industrial regions (Oda & Maksyutov, 2011). Static geographic variables include longitude, latitude, elevation, slope, and aspect.

TCCON observations. TCCON measurements are used only for independent evaluation (Wunch et al., 2011). Site observations are collocated with reconstructed XCO_2 fields in space and time, and only valid site-date pairs during 2018–2020 are included in the site-level comparison. TCCON data are not used for model training.

Preprocessing. Continuous variables are interpolated or aggregated according to their native spatial and temporal resolution, whereas categorical variables are remapped using category-preserving procedures. Dynamic variables are matched to each reconstruction date, while static variables remain fixed. Climate zone, vegetation type, and emission level are additionally used to instantiate the structured process priors described in Section 4.

B Training Objective

GeoPrior-Mamba is trained with sparse OCO-2 supervision at both the 0.25° and 0.05° reconstruction scales. The 0.05° target is obtained from quality-controlled OCO-2 retrievals mapped to the fine grid, while the 0.25° target is obtained by aggregating valid observations to the coarse grid with the corresponding observation mask.

Let y denote the gridded OCO-2 XCO_2 reference field, $\hat{y}$ the model prediction at a given scale, and M the binary observation mask. The masked reconstruction loss is

$$\mathcal{L}_{\mathrm{rec}} = \mathrm{MSE}(M \odot \hat{y},\, M \odot y). \tag{S1}$$

For the 0.05° output, the model additionally predicts a log-variance field s. The uncertainty-aware Gaussian negative log-likelihood is

$$\mathcal{L}_{\mathrm{nll}} = \frac{1}{\sum_i M_i} \sum_i M_i \left[\frac{1}{2} \exp(-s_i)(y_i - \hat{y}_i)^2 + \frac{1}{2} s_i \right]. \tag{S2}$$

The masked MSE anchors point-wise reconstruction accuracy, while the NLL term learns heteroscedastic uncertainty. In implementation, the uncertainty term is introduced with a warm-up schedule to stabilize early-stage training.

The network predicts residual corrections d relative to the large-scale CAMS background x_{bg},

$$\hat{y} = x_{\mathrm{bg}} + d. \tag{S3}$$

Spatial regularization is therefore applied to the residual rather than directly to absolute XCO_2. A Laplacian smoothness term penalizes high-frequency residual artifacts,

$$\mathcal{L}_{\mathrm{lap}} = \left\| \nabla^2 d \right\|_2^2, \tag{S4}$$

with contributions from both the 0.25° and 0.05° residual fields.

To preserve local spatial gradients, let $d_{\mathrm{gt}} = y - x_{\mathrm{bg}}$ denote the target residual. The gradient term is

$$\mathcal{L}_{\mathrm{grad}} = \left\| \nabla_x d - \nabla_x d_{\mathrm{gt}} \right\|_1 + \left\| \nabla_y d - \nabla_y d_{\mathrm{gt}} \right\|_1, \tag{S5}$$

where ∇_x and ∇_y are Sobel gradient operators. This term is applied to the 0.25° residual field over land pixels and is used as a spatial regularizer rather than a sparse supervised loss.

The complete objective is

$$\mathcal{L}_{\mathrm{total}} = \mathcal{L}_{\mathrm{rec}} + \lambda_{\mathrm{nll}} \mathcal{L}_{\mathrm{nll}} + \lambda_{\mathrm{lap}} \mathcal{L}_{\mathrm{lap}} + \lambda_{\mathrm{grad}} \mathcal{L}_{\mathrm{grad}}. \tag{S6}$$

Here, λ_{nll}, λ_{lap}, and λ_{grad} are weighting coefficients. The reconstruction and Laplacian terms include contributions from both output scales, the NLL term is applied to the 0.05° output, and the gradient loss is applied to the 0.25° land residual field.

C Evaluation Protocol

Model performance is evaluated using a date-based split. All valid OCO-2 observations from held-out reconstruction dates are excluded from training and used only for validation or testing. Metrics include MSE, RMSE, MAE, bias, and R^2, computed only at valid OCO-2 observation locations. GeoPrior-Mamba is compared with CAMS background interpolation, Trans-XCO_2, and a knowledge-free Mamba backbone. The reported ablations remove the structured prior branch, the 0.25° branch, or the NLL term. Convergence behavior is assessed from validation trajectories and the number of epochs required to reach a given validation error. Independent TCCON observations provide an additional site-level consistency check.

D Comparison with Previous XCO_2 Reconstruction Studies

Table S2: Comparison with representative XCO_2 reconstruction studies. Reported performance values are taken from the respective studies and are not directly comparable because the reconstruction periods, domains, observations, and validation protocols differ.

Method	Temporal	Spatial	R^2	RMSE	Time frame	Coverage	Literature
DINEOF-BME	Monthly	1°	0.92	1.60	2010–2020	Global (land + ocean)	Gao et al. (2023)
Deep forest	Monthly	0.1°	0.96	0.95	2014–2020	Global land	Zhang et al. (2023)
At-Bi-LSTM	Monthly	0.05°	0.92	1.54	2015–2021	Global land	Wang et al. (2025)
Extreme tree	8-day	0.01°	0.83	1.79	2014–2018	China	Li et al. (2025)
Stacking model	8-day	0.05°	0.98	1.03	2000–2020	Global (land + ocean)	Guan et al. (2024)
Kriging	3-day	1°	0.97	0.88	2009–2020	Global land	Sheng et al. (2023)
M-FRK	Daily	1°	0.93	1.06	2016–2019	Global (land + ocean)	Huang et al. (2024)
Self-supervised	Daily	0.25°	0.93	0.18	2010–2020	Global (land + ocean)	Wang et al. (2023)
DINCAE	Daily	0.1°	0.90	1.44	2018–2022	Global (land + ocean)	Antezana Lopez et al. (2025)
DINEOF-BME	Daily	0.1°	0.99	1.10	2003–2022	Global land	Qu et al. (2026)
U-Net	Daily	0.1°	0.99	1.09	2003–2023	Global land	Wang (2026)
Stacking	Daily	0.1°	0.97	0.97	2016–2020	Global land	Li et al. (2026)
MLE	3-h	0.5°	0.92	2.62	2003–2020	Global (land + ocean)	Jin et al. (2022)
Trans-XCO_2	Daily	0.05°	0.93	1.04	2021–2024	Global (land + ocean)	Meng et al. (2026)
This study	Daily	0.05°	0.93	0.77	2018–2020	Global (land + ocean)	This study

Broader literature context. Table S2 highlights representative products that are close to the present task in resolution, temporal frequency, or coverage. Additional studies considered in the literature survey include satellite-derived and machine-learning XCO_2 mapping over regional and global domains, ensemble and gap-filling approaches, and related carbon-cycle applications (Antezana Lopez et al., 2025; Bhattacharjee & Chen, 2020; Byrne et al., 2023; Chen et al., 2024; Cheng et al., 2022; Friedlingstein et al., 2022; Hammerling et al., 2012; He et al., 2026; Hua et al., 2024; Huang et al., 2024; Hwang et al., 2026; Jin et al., 2022; Li et al., 2026; Liang et al., 2021; Liu et al., 2024; Pan et al., 2021; Qu et al., 2026; Reuter et al., 2020; Sheng et al., 2023; Tang et al., 2024; Tian et al., 2024; Wang, 2026; Wu et al., 2023; Zhang et al., 2023a; Zhang & Liu, 2023). These studies are cited here to keep the main-text related-work discussion concise while preserving a broader record of relevant reconstruction and satellite-CO_2 literature. Because their targets and evaluation protocols differ, they are not treated as directly comparable benchmarks.

The studies in Table S2 span monthly to sub-daily temporal resolution and approximately 1° to 0.01° spatial resolution, with domains ranging from regional or land-only reconstruction to global land–ocean coverage. Because their reported metrics use different periods, observational datasets, and validation strategies, the table should be interpreted as a contextual comparison rather than a direct ranking of methods.

Within this context, GeoPrior-Mamba provides daily global XCO_2 fields at 0.05° resolution over both land and ocean. Trans-XCO_2 is among the most directly comparable products in terms of temporal frequency, spatial resolution, and global coverage. The distinguishing feature of GeoPrior-Mamba is not only its reconstruction resolution, but also its explicit pathway for incorporating structured process context associated with biospheric uptake, ecosystem respiration, and anthropogenic emissions. These priors guide feature learning but do not directly prescribe reconstructed XCO_2 values.