Thanks so much for these thoughtful and constructive comments. We agree that the manuscript could be improved in terms of clarity, specifically regarding the definition of specific knowledge gaps, the establishment of key innovative points, and the integration of our research findings with existing literature. Consequently, during the revision process, we retained the existing model and results framework, focusing instead on enhancing the clarity and coherence of the manuscript. Specifically, this involved more precisely defining the knowledge gap, strengthening comparisons with prior research, clarifying the hierarchical structure of our main contributions, and systematically supplementing the figure captions. Below is our point-by-point response to the revision comments.

Response to comments from the Editor:

1. Introduction and framing of the knowledge gap

The introduction has improved, but the framing of the knowledge gap remains somewhat broad. The manuscript would benefit from a more focused statement of the specific unresolved question being addressed. In particular, please distinguish more clearly between what is already established in prior studies and what this paper uniquely contributes. Related to this, the main novelty should be stated more selectively and explicitly, rather than being distributed across multiple possible contributions.

Reply: Thank you for this important comment and we agree that the knowledge gap was not yet framed with sufficient precision. We have revised the end of the **Introduction** to frame the knowledge gap more specifically and to distinguish more clearly between established understanding and the unique contribution of this study. The objectives and novelty statement have also been revised accordingly. Below are the specific modifications in the tracking version manuscript. The red content is the modified information, and the black content represents the original text (the same applies below).

Line 64-106 in the tracking version manuscript: ~~However,~~ To date, the validation of most Lagrangian-based models, along with their associated simulation efforts, has primarily been conducted at the laboratory or site scale. For regional to basin scale solute transport ~~at the regional scale or basin-scale with~~ processes involving long ~~travel~~migration distances and laterally extensive ~~pathways still lacks~~ flow paths, there remains a lack of quantitative characterization explicitly constrained by hierarchical sedimentary ~~architecture. To date, most Lagrangian-based models' validation and related simulation work have been conducted at laboratory or site scales. Over the past decade, kilometer- to basin-scale investigations have also begun to transition from simplified effective-parameter descriptions toward~~structures. Although recent high-resolution, ~~architecture-resolved analyses. Architectural attributes~~structure-resolving analytical approaches have shown that lithofacies geometry, connectivity, and conductivity ~~(K) statistics have been shown to~~contrasts can organize macroscopic dispersion and non-Fickian ~~behaviour: high-resolution and architecture-resolved models built in the Llagas groundwater subbasin link lithofacies geometry and connectivity to~~behavior at kilometer scale, (Carle et al., 2006; Pauloo et al., 2021; Guo et al., 2019), it remains unclear which hierarchical level of heterogeneity remains transport-dominant under basin-scale flow conditions, how long pre-asymptotic behavior persists, and whether the superposition of numerous large-scale pathways amplifies or buffers predictive uncertainty.

Against this background, this study uses the Qiqihar aquifer system as a basin-scale testbed and combines hierarchical sedimentary architecture reconstruction with groundwater flow and solute transport simulations to investigate the following key questions: (i) Is basin-scale dispersion behavior primarily controlled by dominant large-scale lithofacies structures, or does it still require the explicit characterization of heterogeneity at finer scales? (ii) How do the geometry of solute sources, sedimentary architecture and hydraulic contrasts between lithofacies influence the evolution of macrodispersion and ~~fast pathways/tailing contrasts (Carle et al., 2006); kilometer-scale simulations (Pauloo et al., 2021) demonstrate that aquifer anisotropy and seasonal recharge/pumping-driven shifts in mean flow direction modulate the non-Fickian tails; and regional upscaling research under transient boundaries (Guo et al., 2019) shows that late-time tails cannot be reproduced by the multirate mass transfer (MRMT) model calibrated under steady flow because mass transfer between high- and low-K units~~

~~varies with boundary-driven internal gradients. Even with these advances, architecture-resolved quantification of regional/basin-scale plume evolution under~~its long-term pre-asymptotic behavior? (iii) Under field-representative basin-scale flow conditions, how do structural and hydraulic factors shape transport uncertainty and how uncertainties propagate and evolve? By comparing the results of this study with findings from previous laboratory- and field-scale studies conducted within the same conceptual framework (Dai et al., 2005; Ramanathan et al., 2010; Soltanian et al., ~~heterogeneity and boundary-driven gradients remains limited, particularly regarding the persistence of asymptotic behaviour and its uncertainty~~2015a; Ren et al., 2023), this study clearly distinguishes which controlling factors possess scale-universality and which factors manifest specifically at the basin scale.

~~Accordingly, this study integrates hierarchical sedimentary structures characterization, numerical simulations, uncertainty analysis, using the Qiqihar aquifer system as a basin-scale testbed. Our objectives are to (i) quantify the relative contributions of sedimentary architectural attributes, hydraulic statistics, and source size under field-representative aquifer and large-scale flow fields, (ii) identify which hierarchical scales of heterogeneity dominate solute dispersion, and (iii) evaluate how structural (model-form) or parametric/data uncertainty propagate to transport predictions across scales. This work aims to provide guidance for more reliable predictions of solute transport in field-representative aquifer systems. Furthermore, as mentioned above, existing laboratory and site-scale studies have established mechanistic links between facies geometry/connectivity and solute dispersion (Dai et al., 2005; Ramanathan et al., 2010; Soltanian et al., 2015a). Recent global sensitivity analyses further demonstrate that limited aquifer structure parameters and hydraulic conductivity statistics exert first-order control on kilometer scale dispersion (Ren et al., 2023). These findings enable comprehensive comparisons and discussions at experimental, field, and basin scales within the same research framework, thereby constructing a complete multiscale chain to identify which hierarchical sedimentary architecture and driving factors govern macroscopic spreading and dispersion across scales.~~

2. Discussion and positioning of the main advance

The discussion would be strengthened by more clearly identifying the main new insight from the study. At present, several findings are highlighted, but it is still not fully clear which are intended as the primary advances, which are broadly consistent with prior expectations, and which may be specific to the present modeling setup. Please sharpen the discussion around the actual contributions of the work.

Reply: Thank you for this helpful comment. We have substantially sharpened the **Results** and **Discussion** sections to identify the main new insight more clearly. We have also clarified which findings are consistent with prior expectations and which are more specific to the present basin-scale modeling setup. Below are the specific modifications in the tracking version manuscript.

Line 319-378 in the tracking version manuscript: ~~The close agreement between multiscale and Scale II models, especially under planar-source release (Figure 6c and 6d), demonstrates that at the basin scale, dispersion characteristics can be effectively captured by accurately representing the geometry of controlling lithofacies at larger scales, with diminished influence from smaller-scale heterogeneity. This finding is consistent with the findings of Ramanathan et al. (2010) and Soltanian et al. (2015b) at the site-scale research, who demonstrated similar behaviour using Lagrangian-based macrodispersion models. Comparable patterns were also observed in the site-scale simulations of Ren et al. (2022) at the Borden site, where α(t) exhibited a rapid increase and quickly converged to an asymptotic value.~~

The time derivative of the mass transport distance ~~corresponds to the mean velocity of solute migration in~~ $v_c(t)$ provides a ~~given direction. According to Figure 6a and 6c, the~~simple consistency check on the simulated transport behavior. The average transport velocity of the solute plume is approximately 0.058 m/d for a point-source release and 0.027 m/d for a planar-source release~~. Using the mean hydraulic conductivity of 12.05 m/d (Table 2), a regional hydraulic gradient of ~1‰, and an effective porosity of ~0.35,~~ (Figures 6a and 6c). The latter is close to the Darcy-based estimate of the mean groundwater velocity~~ is~~, about 0.03 m/d~~. The consistency between this .~~, calculated ~~velocity and the simulated plume migration suggests that the model~~

~~adequately captures~~from the mean hydraulic conductivity of 12.05 m/d (Table 2), regional hydraulic gradient of ~1‰, and effective porosity of ~0.35 used in this study. This consistency indicates that the flow and transport simulations reproduce the macroscopic solute ~~transport characteristics in the study area. Comparisons with the Borden site highlight distinct timescales: $\alpha(t)$ at Borden increased rapidly and reached its asymptotic value within a relatively short time, whereas in this~~ migration behavior of the study ~~$\alpha(t)$ approached its asymptote more gradually (Ren et al., 2022).~~ system.

A first-order result is the close agreement between the multiscale and Scale II simulations, particularly under planar-source release. This ~~difference likely reflects the higher mean velocity at the Borden site (~0.091 m/d), which accelerates stabilization~~indicates that basin-scale plume evolution can be captured primarily by the geometry and connectivity of the dominant large-scale lithofacies, whereas finer-scale heterogeneity mainly modulates the details of transport. The source geometry mainly affects early-time sampling of heterogeneity and thus the magnitude of realization-to-realization variability. Under point-source release, the plume initially samples only a limited portion of the K field, resulting in larger fluctuations in transport distance and solute dispersion. In contrast, the planar source intersects a broader set of flow paths from the outset, which narrows the uncertainty envelopes and yields a more stable ensemble-like response. This result is consistent with previous theoretical and numerical studies showing that larger source areas enhance early sampling of heterogeneity, stabilize large-scale transport statistics, and reduce uncertainty in plume behavior (Dagan, 2017; de Barros, 2018). This also confirms that source dimensions remain important even at the basin scale.

~~At the release location, high proportions of gravel and medium-to-coarse sand lithofacies lead to strong local-scale heterogeneity in the K field. Under point source release scenario, the plume initially samples a limited portion of the heterogeneity, local scale flow velocity deviate significantly from average regional groundwater flow velocity, showing a large fluctuation amplitude and significant uncertainty (shaded areas in Figures 6a and 6b). In contrast, planar source covers a broader portion of the heterogeneous medium at early times. This "source-area enlargement effect" smooths~~

~~local-scale velocity deviations and explains why confidence intervals are much narrower in the planar-source scenarios. Similar trends were reported by Cao et al. (2018) and de Barros (2018), who showed that enlarging the source area/width markedly narrows uncertainty bands and reduces realization-to-realization spread of plume metrics in heterogeneous aquifers. Theoretically, Dagan (2017) has explained why planar, large-area injections approach ergodic sampling of the conductivity field, by smoothing local velocity fluctuations and stabilizing large-scale transport statistics.~~

~~The ensemble macro-dispersivity ($\alpha_{ens}(t)$) is more commonly used due to its simpler definition. Moreover, $\alpha_{ens}(t)$ often overestimates the solute dispersion under non-ergodic or pre-asymptotic conditions. Figure 6b illustrates this overestimation significantly and shows that the error associated with $\alpha_{ens}(t)$ increases progressively with transport time for both the multiscale model and the larger-scale model. The persistence of this overestimation until the end of the 10000-day simulation suggests that plume evolution remains largely governed by advection and has not yet reached a dispersion-dominated regime controlled by local-scale dispersion. Once the solute plume samples a sufficiently large domain of the heterogeneous medium, realization variability is substantially reduced and large-scale transport properties converge to the ensemble mean. Accordingly, $\alpha_{eff}(t)$ tends toward $\alpha_{ens}(t)$, consistent with theoretical predictions.~~

Figure 6 also shows that α_{eff}(t) and α_{ens}(t) do not fully converge within the 10000-day simulation period, although the degree of separation varies with source geometry and transport stage. In general, the difference between the two metrics remains more evident under point-source release condition. This incomplete convergence indicates that plume evolution remains in a prolonged pre-asymptotic state under the simulated basin-scale transport conditions. At this stage, the plume's evolution has not yet transitioned into the dispersion-dominated phase, which is characterized by the dominance of local-scale dispersion processes. Compared with previously reported site-scale results at the Borden site, where macrodispersivity approached a quasi-stationary value over a much shorter interval, the present basin-scale system exhibits a substantially slower approach to convergence.

Line 400-415 in the tracking version manuscript: Figure 7 shows the simulation results for different scenarios. With the increase of time, the solute transport distance increases proportionally, ~~and~~while the $\alpha_{eff}(t)$ shows a power function growth trend and approaches a quasi-stable value after roughly 5000 days. Across the tested scenarios, the asymptotic-scale $\alpha_{eff}(t)$ remains on the order of 170 m, indicating that the basin-scale system remains strongly dispersive over long travel times. This behaviour is consistent with the well-established theory of large-scale dispersion in heterogeneous media, where transport is characterized by an initial non-Fickian regime followed by a transition towards Fickian behaviour at late times (Dagan 1989; Neuman and Tartakovsky, 2009). ~~Across~~Increasing the ~~tested scenarios, the $\alpha_{eff}(t)$ stabilizes around 170 m after approximately 5000 days. High permeability lithofacies type (~~proportion of GCS~~) always provides preferential pathways that accelerate~~ enhances plume migration~~, while low permeability type (SC) acts as barriers that restrict~~ and leads to larger $y_c(t)$ and $\alpha_{eff}(t)$ values, whereas increasing the proportion of SC suppresses both plume ~~spreading and extent solute residence times. This demonstrates that even at the basin scale,~~ migration and dispersivity. These trends indicate that modifying lithofacies proportions ~~can still reshape~~changes the balance between connected preferential pathways and retarding ~~controls. Both high~~low-*K* ~~and low-*K* lithofacies exert strong controls on plume transport behaviours~~domains, even though the overall basin-scale response remains relatively smooth. Previous theoretical and numerical studies (Fiori et al., 2010; Amooie et al., 2017; Soltanian and Ritzi, 2014; Puyguiraud et al., 2020), have also confirmed that large-scale solute dispersion is governed not by a single conductivity class, but by the combined influence of extreme permeability contrasts.

Line 421-435 in the tracking version manuscript: ~~A notable finding is that even when the proportion of GCS increases, the differences in plume metrics among scenarios remain modest, and the~~ The uncertainty ~~bands widen only slightly. Actually, the plume travels only a short distance after 10000 days for this basin-scale system (~~envelopes in Figure ~~7a), indicating that non-ergodic conditions persist for extremely long times. This suggests that, changes in lithofacies proportions adjust transport characteristics,~~7 also show a systematic but ~~the overall plume dynamics are buffered by the immensity of the~~

~~flow domain. At regional basin scales, modifications~~moderate response to lithofacies ~~proportions may act as secondary drivers of plume extent relative to other large-scale controls.~~

~~Realization-to-realization variability also reveals a clear scale-dependent trend.~~proportion. Increasing the ~~proportion~~fraction of high-*K* ~~lithofacies enhances solute mobility but simultaneously~~facies generally amplifies ~~realization-to-realization~~the variability among realizations, whereas increasing the fraction of low-*K* ~~lithofacies reduces both transport velocity and uncertainty, yielding~~facies produces more uniform outcomes. However, ~~uncertainties do not grow indefinitely with conductive lithofacies dominance but instead plateau due to kilometer-scale~~ the differences among scenarios remain much smaller than would typically be expected at the site scale. This suggests that, under basin-scale flow conditions, the effect of composition changes is present but moderated by large-scale spatial averaging.~~ Such findings highlight the importance of explicitly accounting for scale-dependent controls when extrapolating transport models from site to regional contexts, especially for risk assessment and large-scale groundwater management.~~

Line 455-497 in the tracking version manuscript: Figures 8a-8c show the results for solute transport distance, and Figures 8d-8f are the $\alpha_{eff}(t)$ results. They all show a clear, asymmetric response when the mean *K* is tripled for a single lithofacies while keeping variance and architecture fixed. Increasing the mean *K* of the most permeable lithofacies (GCS) accelerates plume migration, raises the asymptotic α_{eff}~~, and slightly widens the uncertainty bands. Conversely, tripling the mean *K* of the medium-permeability lithofacies (MFS) counterintuitively reduces migration distance and affects the asymptotic convergence of dispersion, whereas perturbing the *K* value of SC produces virtually no change. By calculating the global mean *K* of the 50 realizations, it can be obtained that the global mean *K* value changes from 12.05 m/d of Qiqihar site to 22.20 m/d, 19.49m/d and 14.73 m/d after expanding the mean *K* values of the three lithofacies. Logically, MFS perturbation should have facilitated solute transport. Further analysis of the constructed heterogeneous sedimentary architecture model reveals that the dominant lithofacies progressively transition from gravelly/medium-~~

~~coarse sand to medium-fine sand from upstream to downstream. MFS lithofacies is less distributed at the solute plume release location. As the conductivity in the model is generated based on lithofacies distribution, this may be a major reason for the anomalies in the transport distance results. Temporal variations in $\alpha_{eff}(t)$ also highlight the contribution of individual facies in shaping plume evolution. In the GCS perturbation, $\alpha_{eff}(t)$ increases sharply from the onset of solute release, and its asymptotic value remains well above the baseline case, emphasizing the dominance of coarse type connectivity in channelling solute migration. In the MFS perturbation, $\alpha_{eff}(t)$ decreases in the early stages but then gradually converges to asymptotic equilibrium. This pattern reflects spatial averaging at the basin scale eventually restores ensemble-like behaviour. These early and late dynamics are~~(t), and slightly widens the uncertainty bands. This indicates that the most permeable lithofacies has a strong influence on basin-scale plume migration. By contrast, increasing the mean K of MFS does not accelerate plume migration, even though the domain-scale mean conductivity increases (from 12.05 m/d of Qiqihar site to 19.49m/d). Instead, the transport distance is reduced relative to the baseline case, and the corresponding α_{eff} curve shows a different evolution toward quasi-stationary behavior. Increasing the mean K of SC produces only a very limited change in both transport distance and $\alpha_{eff}(t)$. The early and late dynamics of $\alpha_{eff}(t)$ are also consistent with observations of non-Fickian transport, where temporary suppression and delayed convergence are characteristic of strongly heterogeneous aquifers (Fiori and Dagan, 2000; Dentz et al., 2011). However, more importantly, the spatial averaging effect at the basin scale ultimately restores ensemble-like behaviour.

~~Collectively, the results indicate that macro-dispersion may not simply controlled by the average conductivity of the system, but by the extent to which variations in facies properties reallocate groundwater fluxes across the underlying connected pathways. Specifically, enhancing GCS strengthens pre-existing preferential pathways, accelerating~~The contrasting responses are likely related to the spatial roles of the three lithofacies in the modelled architecture. GCS is the most conductive lithofacies and is more influential near the release area, so increasing its mean K reinforces pre-existing fast routes and directly promotes plume migration. In contrast, ~~increasing~~ MFS is less dominant at the ~~mean *K* of MFS~~ release location and becomes more important farther along the transport domain. The mean *K* perturbation therefore reduces the contrast with GCS~~,~~ and redistributes part of the groundwater flux ~~into slower pathways, and thus~~ away from the fastest paths, which suppresses early-time ~~dispersion. This complex flux redistribution mechanism effectively suppresses the overall solute migration rate. For SC, even~~plume spreading. Even after a threefold increase~~ leaves it far~~, SC remains much less conductive than the other ~~facies, preventing it from contributing~~two lithofacies and therefore cannot contribute substantially to the connected high-*K* transport network. ~~Such asymmetric responses are consistent~~These results show that basin-scale transport cannot be interpreted solely from the change in domain-average conductivity, because the effects of *K* perturbation depend strongly on which lithofacies is modified and where that lithofacies occurs within the connected architecture. These findings align with previous theoretical and numerical studies, which ~~demonstrate~~indicate that solute dispersion at large scales emerges from the interplay between ~~facies~~lithofacies contrasts and connectivity, rather than from the mean conductivity values of single ~~facies~~lithofacies (Zinn and Harvey, 2003; Soltanian and Ritzi, 2014). ~~All these observations suggest that model calibration that focus only on asymptotic metrics may overlook early-time transport features that are critical for risk assessment and monitoring system design.~~

3. Engagement with prior literature

Both the introduction and discussion would benefit from somewhat stronger connection to prior literature. In particular, I encourage the authors to compare their interpretations more explicitly with earlier studies on regional- and basin-scale transport, architecture-resolved modeling, and related non-Fickian transport behavior. This would help clarify the paper’s novelty and significance.

Reply: Thank you for this constructive suggestion. We have strengthened the connection to prior literature in both the **Introduction** and the **Discussion** by engaging more explicitly with previous studies on regional- and basin-scale transport, architecture-resolved modeling, and persistent non-Fickian behavior. These revisions help position the novelty and significance of the present study more clearly. The revision for the Introduction section can be found in the response to comment 1. Below are the specific modifications in the **Discussion** section.

Line 499-603 in the tracking version manuscript: ~~Basin-scale simulations indicate that conservative solute dispersion and its uncertainty are principally organized by facies conductivity contrasts and the geometry of the architectural elements at the larger scale, whereas modest perturbations to smaller-scale structure produce comparatively minor changes in transport metrics. This is consistent with the multiscale sedimentary-architecture based theoretical analysis and with prior site-scale simulation studies. This study also distinguishes between $\alpha_{eff}(t)$ and $\alpha_{ens}(t)$: $\alpha_{ens}(t)$ increasingly overestimates the time-varying dispersion, while $\alpha_{eff}(t)$ measures spreading within a typical realization and approaches a stable value as sampling increases. The kilometer-scale $\alpha_{eff}(t)$ obtained in this study fall within the classical field scale statistics compiled by Gelhar et al. (1992), who found $\alpha(t)$ typically in the tens to hundreds of meters at similar scales.~~

~~A central finding is the scale-dependent nature of plume stabilization. Over a 10000 days simulation period, $\alpha_{eff}(t)$ initially increased slowly and then stabilized, but this stabilization occurs over a relatively prolonged period (around 5000 days), indicating that plume evolution remains in a pre-asymptotic (quasi-ergodic) state for thousands of days at the basin scale. This behaviour contrasts with site-scale results at the Borden site (Ren et al., 2022), where $\alpha_{eff}(t)$ was shown to reach an asymptotic value much more rapidly (after 400 days), plausibly facilitated by higher mean groundwater velocities. In the present basin-scale system, the solute explores only a small fraction of the heterogeneous flow field over a considerable period of time, the effective transport response is "buffered" by the domain immensity. Source size exerts a qualitatively consistent effect across scales: expanding the source promotes broader early-time sampling of heterogeneity and thereby reduces inter-realization variability, yielding narrower uncertainty during transport process. By contrast, the uncertainty band for $\alpha(t)$ tends to stabilize or even decrease later at the basin scale, again reflecting domain scale averaging that mutes source geometry effects far downstream. In the site-scale tracer experiments (e.g., MADE, Borden), meter-scale K variability may lead to~~

~~apparent velocities differ by orders of magnitude between observation points, and strongly connected preferential pathways commonly yield pronounced early arrivals and heavy-tailed breakthrough curves (Zheng et al., 2011; Bianchi and Pedretti, 2017), thereby increasing predictive uncertainty in plume evolution and arrival-time statistics. Simulations at the Borden site well demonstrate that the increased proportion of more permeable lithofacies significantly amplifies solute dispersion and output uncertainty. Basin-scale models, however, exhibit a weaker response, consistent with the buffering effect of long travel distance and multiple overlapping pathways.~~

~~This study further highlights that even at the basin scale, $\alpha(t)$ is not a simple constant. It is also closely related to the aquifer's heterogeneous structure and its hydraulic parameters. Scenario analysis reveals that the facies proportion and mean-conductivity perturbations illustrate two complementary mechanisms by which multiscale heterogeneity governs solute dispersion. Perturbations in facies proportions alter the statistical balance between preferential pathways and retarding domains, whereas perturbations of mean conductivity reshape the contrast among facies types, redistributing fluxes across the connected network. At the basin scale, both mechanisms exert measurable influences on plume dynamics, but the divergence among scenarios remains modest due to kilometer-scale spatial averaging and persistent non-ergodic conditions. Recent~~The present results clarify a central issue raised in the Introduction: under basin-scale flow conditions, the dominant controls on conservative solute dispersion are the geometry and connectivity of the larger-scale lithofacies framework, whereas finer-scale heterogeneity mainly modulates the details of plume evolution. Viewed together, the scenario analyses show that lithofacies proportions and mean K are not separate or competing explanations for this behavior, but two complementary ways in which the same architectural framework governs transport. Changing lithofacies proportions alters the statistical balance between preferential pathways and retarding domains, whereas changing the mean K of individual facies modifies how groundwater flux is partitioned among those pathways. What emerges from these results is that basin-scale dispersion is controlled less by domain-average hydraulic properties than by the way conductivity contrasts are embedded within the connected lithofacies architecture. This interpretation is consistent with previous regional and basin scale studies showing that aquifer-system heterogeneity, lithofacies architecture, and pathway organization remain influential controls on large-scale plume evolution (Carle et al., 2006; Pauloo et al., 2021). A recent global sensitivity analysis across multiscale heterogeneous media shows a robust ranking for non-reactive solute dispersion~~:~~ (Ren et al., 2023): the ~~facies~~lithofacies mean *K* is typically the most influential factor, followed by ~~facies~~ volume proportions and ~~facies~~ mean lengths~~;~~, whereas variance and some correlation scales contribute less. ~~Importantly, when the heterogeneity integral scale reaches 100m to 1000m, the regional hydraulic gradient becomes non-negligible for non-reactive transport. Our~~These basin-scale results are also broadly consistent with this ranking, especially in showing that lithofacies mean *K* and proportions exert the clearest influence on basin-scale dispersive behavior.

This study also indicates a markedly delayed approach to quasi-steady behavior at the basin scale. $\alpha_{eff}(t)$ approached a quasi-steady state only after roughly 5000 days, and $\alpha_{eff}(t)$ and $\alpha_{ens}(t)$ did not fully converge within the entire 10000-day simulation period. This persistent separation indicates that plume evolution remains strongly pre-asymptotic over long times and distances. Compared with the Borden tracer test study, where dispersivity approached quasi-stationary behavior after around 400 days (Ren et al., 2022), the present basin-scale system shows a substantially slower approach to convergence. Moreover, site-scale studies have also shown stronger sensitivity of plume behavior and predictive uncertainty to realization-specific pathways and source sampling (Zheng et al., 2011; de Barros and Dentz, 2016), whereas the basin-scale response here is more muted. These differences reflect not only domain-scale spatial averaging, but also the gradual way in which the plume samples the heterogeneous flow field and the delaying effect of multiple overlapping pathways on the emergence of ensemble-like transport. Within this setting, source geometry mainly influences early-time uncertainty by determining how much heterogeneity is sampled at the outset, while the long-term structure of plume evolution remains organized by the dominant large-scale architecture. These interpretations are also consistent with previous regional and basin scale studies showing that heterogeneous flow systems can sustain persistent non-

Fickian transport over long distances and times, with a delayed approach to simplified or asymptotic transport behavior (Guo et al., 2019; Pauloo et al., 2021).

Combined with earlier laboratory sandbox experiments and site-scale studies developed within the same hierarchical framework (Ma et al., 2022, 2025), the present basin-scale results help refine how the role of heterogeneity should be understood across sedimentary scales. At local and site scales, conductivity contrasts, lithofacies geometry, and connectivity are often expressed more directly through realization-sensitive plume spreading, stronger source-size dependence, and more immediate pathway control on uncertainty (Dai et al., 2004; Ramanathan et al., 2010; Yin et al., 2020). At the basin scale, however, these ~~rankings~~same controls remain operative, but their expression is increasingly filtered by long travel distances, pathway superposition, and spatial averaging, so that plume evolution becomes more slowly convergent and less sensitive to fine-scale structural differences. In this sense, the controlling mechanisms are not replaced at larger scales, rather, they are reorganized through scale-dependent averaging. In short, sedimentary heterogeneity exerts a scale-persistent but scale-reorganized control on solute transport.

This broader view also holds practical implications for model construction and interpretation. For problems focusing on bulk plume transport at regional to basin scales, it is not necessarily required to retain every detail of the fine-scale stratigraphic structure, provided that the dominant geological contrasts, lithofacies proportions, and directional organization of the connected flow framework are retained. This does not imply that finer-scale heterogeneity is irrelevant, on the contrary, it may remain important for early-time transport, near-source prediction, and problems sensitive to local concentration gradients. It must be acknowledged that neglecting porosity variations and molecular diffusion processes in this study may lead to an underestimation of early plume smoothing and lateral mixing, potentially delaying a significant convergence to Fick behaviour. However, at the basin scale and in the long-distance travel considered in this paper, structure-controlled velocity variations are expected to dominate the dispersion ~~index;~~response, therefore, the main conclusions regarding ~~relative~~the roles of lithofacies proportions, conductivity contrasts, and

~~connectivity~~connected architecture remain unchanged. In that sense, the findings of this study support a scale-aware modeling strategy, in which the level of model complexity should be matched to the transport behavior of interest, rather than being increased indiscriminately across all scales. Such a scale-aware perspective provides a more practical theoretical basis for model simplification, monitoring design, and uncertainty assessment in data-limited regional aquifer systems.

~~Independent 3-D tank/column experiments conducted with sediments from the same Nen River setting (Ma et al., 2022, 2025) provide small-scale mechanistic insight into the scaling of $\alpha(t)$. By explicitly tying the growth of $\alpha(t)$ to velocity contrasts caused by K variability, to lithofacies geometry, and to cross-facies transition, a pronounced scale dependence was observed: heterogeneous mixtures produced a much steeper growth of longitudinal $\alpha(t)$ than single-facies columns, with best-fit longitudinal $\alpha(t)$ on the order of 0.1 m. The consistency between the experimental evidence from this place and basin-scale model results supports our central conclusion that K-contrast and connectivity-governed mechanisms remain influential across multiple scales.~~

~~Taken together, the evidence from our study and a review of the broader literature supports three key points: (i) Solute dispersion and plume evolution are consistently shaped by coupled contrast- and connectivity-driven processes across scales; (ii) the apparent influence of these mechanisms is scale dependent and prolonged pre-asymptotic behaviour and pathway superposition buffer bulk transport responses; and (iii) it is sufficient to preserve dominant geological contrasts, lithofacies proportions, and directional mean lengths for regional-scale prediction, while problems near a release source or at early transport stage still require finer architectural resolution and careful treatment of source size. More broadly, this study emphasizes that reliable regional forecasts require not only calibrating mean hydraulic properties but also retaining the geological contrasts and connectivity that govern transport, together with explicit recognition of scale-dependent averaging effects. These insights advance a more integrated understanding of basin-scale solute transport process and provide a conceptual basis for monitoring and management designs that remain robust under limited site characterization and persistent predictive uncertainty.~~

Response to comments from Anonymous Referee #1:

1. Line 75: I suggest expanding the summary of prior research to highlight the most relevant key findings more explicitly. This would help sharpen the scientific context and clarify the specific research question, which currently remains somewhat broad.

Reply: Thank you for this helpful comment. We have revised the **Introduction** to sharpen the scientific context and to summarize the most relevant prior findings more explicitly. In particular, we now distinguish more clearly between what has already been established at laboratory/site and regional/basin scales and the specific unresolved questions addressed in this study. The revision for the Introduction section can be found in the response to comment 1 from the Editor.

2. Figure captions: Most figure captions are still quite minimal and do not provide enough information for the figures to be fully understood on their own. As a general rule, each figure and its caption should be able to stand alone without requiring the reader to consult the main text.

Reply: We appreciate this comment. We have revised the figure captions to improve their stand-alone readability, especially for the main result figures, by clarifying the plotted quantities, source-release conditions, color coding, and uncertainty envelopes. Below are the specific modifications in the tracking version manuscript.

Line 227-229 in the tracking version manuscript: **Figure 3. An example of multiscale model: (a) 3-D facies, (b) 2-D sections; and Scale II model: (c) 3-D facies, (d) 2-D sections. Section lines are located at *x* = 600m, 3000m, 7600m, 11000 m and *y* = 8000m, 12000m, 16000m, 18000m**

Line 244-246 in the tracking version manuscript: **Figure 4. Conceptual hydrogeologic model of the study area and schematic diagram of solute plume planar source release conditions. The left side corresponds to the Nen River, and the pollution source is oriented approximately perpendicular to the regional groundwater flow direction.**

Line 327-330 in the tracking version manuscript: **Figure 6. ~~Results of~~Temporal evolution of longitudinal (a, c) solute transport ~~distances~~distance $y_c(t)$ and (b, d) macro-dispersivity ~~on two scales with time under two release conditions~~$\alpha(t)$ for multiscale (red)**

and Scale II (blue) models under point-source release (a, b) and planar-source release (c, d) conditions. Solid lines denote ensemble means, and shaded envelopes represent the 10th–90th percentile ranges.

Line 417-420 in the tracking version manuscript: Figure 7. ~~Results~~Effect of lithofacies proportions on longitudinal (a) solute transport ~~distances~~distance $y_c(t)$ and (b) macro-dispersivity $\alpha(t)$ under ~~different lithofacies volume proportion scenarios~~planar-source release condition. Group A and Group B represent more mixed and fine-dominated architectures, respectively. Solid lines denote ensemble means, and shaded envelopes represent the 10th–90th percentile ranges.

Line 449-454 in the tracking version manuscript: Figure 8. ~~Results of (a) solute transport distances and (b) macro-dispersivity under the variations in lithofacies mean hydraulic conductivity~~Effect of the mean K of individual Scale II lithofacies on longitudinal (a) solute transport distance $y_c(t)$ and (b) macro-dispersivity $\alpha(t)$ under planar-source release condition. In each scenario, the mean K is increased threefold while the lithofacies architecture and inter-facies variance remain unchanged. Solid lines denote ensemble means, and shaded envelopes represent the 10th–90th percentile ranges.