Highlights

STAMNet - A Spatiotemporal Attention Module and Network for Upscaling Reactive Transport Simulations of the Hyporheic Zone

Marc Berghouse, Rishi Parashar

- STAMNet allows for rapid upscaling of reactive transport simulations of bioremediation in the hyporheic zone.
- The spatiotemporal attention module (STAM) uses cross-dimensional relationships to extract complex features and is shown to improve performance more than efficient channel attention methods.
- STAMNet shows strong performance for upscaling of biomass, electron donor, and chromium concentrations.
- STAMNet-Upsample uses the STAM architecture to rapidly increase simulation resolution with high accuracy.

STAMNet - A Spatiotemporal Attention Module and Network for Upscaling Reactive Transport Simulations of the Hyporheic Zone

Marc Berghouse[a,b], Rishi Parashar[b]

[a] *Graduate Program of Hydrologic Sciences, University of Nevada, Reno*
[b] *Division of Hydrologic Sciences, Desert Research Institute, Reno*

Abstract

Reactive transport (RT) simulations are important tools for understanding and predicting phenomena in the subsurface. However, RT is computationally intensive and complex simulations can be numerically unstable. Here, we present STAMNet, a low-parameter attention-based suite of neural nets that can upscale and upsample reactive transport simulations, applied to example problem of bioremediation in the hyporheic zone. We show that a simple MLP offers 30x speedup over standard multiphysics RT simulations and can accurately ($\approx 90\%$ R^2) predict the output of multiple variables of a 1x20 meter RT simulation by using the output from a 1x2 meter simulation as input. We add efficient channel attention to our optimized MLP which significantly improves the mean average error but doesn't affect the R^2. We further develop a novel spatiotemporal attention module (STAM), which results in improvements both in mean square error and R^2 (92.5%). Finally, we present a network architecture that utilizes STAM to accurately (99.9% R^2) upsample simulations in two dimensions. Specifically, our model allows for the 2x upsampling of simulations in the x and y dimensions to convert a coarse-grained input into a fine-grained output. These models have potential use for Monte-Carlo-style investigations of bioremediation and the work presented serves as a proof-of-concept for accurate prediction of large sets of spatiotemporal outputs.

Keywords: Reactive Transport, Deep Learning, Attention, Upscaling, Hyporheic Zone

---

1. Introduction

In the vast realm of environmental science, the hyporheic zone (HZ) stands out as a complex interface that has captured the attention of researchers for decades [1, 2, 3]. This subsurface region, generally defined as the interface between river water and groundwater, hosts a myriad of complex interactions [4], with biofilms serving as a central character influencing broader hydrological and geochemical cycles [5].

Multiphysics simulators that use analytical and numerical methods to solve systems of equations that describe hydro-biogeochemical interactions in the subsurface environments, also known as reactive transport (RT) simulators, such as PFLOTRAN [6], STOMP [7], and CrunchFlow [8], are generally considered the gold standard for simulations of phenomena in the HZ. However, large scale RT simulations, and Monte-Carlo-type investigations of RT simulations, have high computational complexity and cost which are sensitive to convergence criteria [9, 10] causing numerical instability and challenges in supporting hydro-biogeochemical research efforts.

Several recent studies in the field of computer science have shown that accurate multi-physics simulation emulation is possible with deep learning [11, 12, 13, 14, 15]. Thus, to alleviate the common shortcomings of RT simulators, some studies have attempted to apply these emulation frameworks to RT data [16, 17, 18, 19, 20, 21]. Laloy and Jacques presented some of the earliest studies that looked into RT emulation with deep learning. They found deep neural networks (DNNs) outperform polynomial chaos expansion networks for the prediction of a target RT variable given some input variables of the RT timeseries. Although emulation is still a popular topic, much of the current research in this domain also seeks to upscale micro and pore-scale models to the macro/continuum scale. Wang and Battiato (2024) provide a comprehensive framework to upscale RT in fracture-matrix systems. Their framework uses a combination of traditional RT algorithms with a recurrent neural network (RNN) to capture the impact of small-scale features, which they show results in improved accuracy compared to a pure macroscale model. You et al. (2024) used convolutional neural networks (CNNs) to upscale pore-scale simulations to continuum-scale simulations. They found that the effective surface area and effective diffusion coefficient could be predicted with high accuracy, but permeability is difficult to predict. These frameworks represent significant advances in the field of RT modeling, although they suffer from a lack of easy integration with current popular methods, and are constrained in their scope. General models that can be easily implemented would increase access of reactive transport simulation tools to a larger community of researchers.

In this paper, we present STAMNet, a deep-learning-based method for the upscaling and upsampling of RT simulations of biomass growth in the HZ. We chose to model biomass growth in the hyporheic zone due to the complexity of the simulations and its importance for many biogeochemical functions they serve. Furthermore, simulations of biomass growth result in outputs at large scales that could be very different than outputs at small scales, which necessitates a more careful upscaling than a simple interpolation or polynomial fit. Furthermore, we consider our simulations of biomass growth in the hyporheic zone to be a proxy for general reactive transport modeling, since the biomass growth simulations take advantage of most of the modeling capabilities in PFLOTRAN. In addition to biomass growth, we use STAMNet to upscale chromium and molasses, highlighting our framework's overall capabilities for the upscaling of bioremediation simulations. We test the performance of an optimized MLP, the MLP + efficient channel attention, and the MLP + our spatiotemporal attention module, STAM, which we find to generally outperform the other models. We test our upscaling method to predict the spatiotemporal output for a 1x20 meter simulation given a 1x2 meter simulation as input. This model allows for a 30x speedup in the generation of large-scale simulations with an R^2 of the predicted mean time series of 92.5%. We also devise an optimized linear architecture for the task of upsampling, which takes a 1x2 meter simulation with a resolution of 100 voxels/m as input and outputs a 1x2 meter simulation with a resolution of 400 voxels/m.

2. Methods

This study uses multiphysics simulations to explore biomass growth in the HZ, and deep learning models to upscale and upsample these simulations. In this section, we describe the boundary conditions and parameters used for our simulations, and the model architectures and training/testing procedures used for our upscaling and upsampling frameworks.

2.1. Simulations of the Hyporheic Zone

2.1.1. General Description of Simulations

Our simulations are based in PFLOTRAN, a multi-physics reactive transport simulator developed by multiple national laboratories [6]. PFLOTRAN represents a state-of-the-art computational framework for simulating coupled subsurface flow and reactive transport processes across multiple spatial and temporal scales. This massively parallel reactive transport code integrates sophisticated numerical methods to resolve multi-phase and hydro-biogeochemical interactions. The code's architecture enables the simulation of various subsurface processes, including density-dependent flow, variable saturation conditions, and non-isothermal phenomena, alongside comprehensive biogeochemical reactions such as aqueous complexation, mineral precipitation/dissolution kinetics, surface complexation, ion exchange, and microbially-mediated transformations. As discussed in the introduction, we seek to use this reactive transport simulator to model biomass growth in the hyporheic zone. To this end, we have specifically adapted the Chrotran [22] version of PFLOTRAN to represent bioremediation in the hyporheic zone at the Darcy scale. Chrotran defines biomass growth as a function of electron donor (ED) concentration through simple Monod kinetics. It uses biotic and abiotic reactions to model Cr(VI) reduction, defines a mobile-immobile mass transfer system for biomass and ED, and allows for bioclogging modeling capabilities via the dependence of porosity and permeability on biomass concentration. For a full description of the biomass growth model, please refer to the original Chrotran paper.

The simulations described in this paper were created for the purpose of modeling certain interactions in the hyporheic zone. We simulate different flow conditions, permeability conditions, and concentration inputs to train our models on general representation of bioremediation simulations in the hyporheic zone. The simulations are not resource limited – simulations with high concentration of nutrients allowed for relatively linear biomass growth throughout the time frame of the simulations (up to 228 days) whereas biomass growth leveled off more towards the end of the simulations for scenarios with low concentration of nutrients,. We chose to not investigate nutrient-limited scenarios because we observed less differences between small scale and large scale simulations in cases of nutrient limitation, meaning a model that allows mapping between the two would be less useful. Furthermore, the primary purpose of our upscaling model (STAMNet) is to provide a means for rapid generation of bioremediation simulations for a variety of input conditions. In most cases of bioremediation, biomass growth is stimulated through the injection of nutrients (i.e., an electron donor), so our focus on high-nutrient simulations enhances the model's applicability to bioremediation.

2.1.2. Boundary Conditions

For the baseline, we simulate a 1 meter (in vertical, or direction of hyporheic flow) by 2 meter longitudinal (in direction of river/groundwater flow) slice of a synthetic hyporheic zone represented by 100 by 200 voxels (dx = dy = 0.01 m). The top and bottom boundaries (1 m difference) respectively represent the surface and bottom-HZ pressures (which controls the amount and direction of vertical flow), and the left and right boundaries (2m difference) represent the pressure gradient in the longitudinal direction, thus controlling the vector of groundwater flow (also referred to here as the horizontal flow). The horizontal pressure gradient is constant over the duration of any given simulation, and the vertical pressure gradient for any given simulation is derived from three different sets of in-situ hyporheic flux data [23, 24, 25] (Fig. 1). The "10x scale simulation" represents a 1 meter by 20 meter slice of the hyporheic zone with the same resolution as the baseline simulations. Importantly, we note here that the pressure boundary conditions are the same for each pair of baseline and 10x simulations. This means the pressure gradient changes between a baseline simulation and its respective 10x simulation. This constant pressure boundary framework is useful for extending simulations of bioremediation in which a specified pressure is established due to stimulant injection, although we also recognize that an upscaling method for constant pressure gradient would be beneficial for general studies of the hyporheic zone.

As discussed in further sections of the methods, all simulation variables, including the horizontal and vertical pressure gradients, take on different values for different simulations. From these base time series, we introduce random variations (large variations for the "high speed" time series and small variations for the other time series in Figure 1) to increase the variability in potential flow conditions for our models to be trained on. The base set of pressure gradients for the horizontal flow was determined a range of realistic groundwater flow rates. Both horizontal flow (V_y) and vertical flow (V_x), and transport, are regulated by Dirichlet boundary conditions.

The primary motivation of this study is to develop upscaling and upsam-

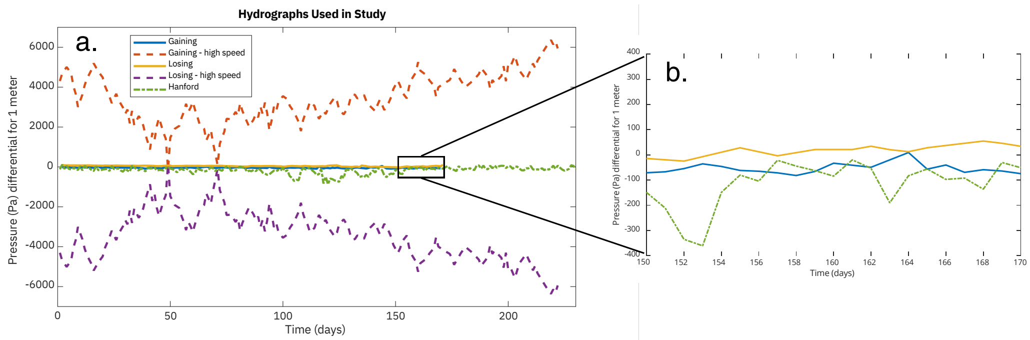

Figure 1: Different flow boundary conditions used in this study. Pressure time series were calculated from in-situ measurements of hyporheic flux ([23, 24, 25]). The gaining boundary conditions represent flux from the groundwater to the surface water, and losing boundary conditions represent flux from the surface water to the groundwater.

pling methods for RT simulations (specifically in the context of bioremediation) using deep learning (Fig. 2). Thus, we generated pairs of simulations for training that are identical in every way except scale (for upscaling) or resolution (for upsampling). As discussed above, the baseline simulations represent a 1 meter by 2 meter slice of the HZ. For the upscaling task, all models use the baseline 1x2 meter simulation as input to predict a 1x20 meter simulation. For the upsampling task, all models use the baseline 1x2 meter simulation with $dx = dy = 0.01$ m as input to generate a 1x2 meter simulation with $dx = dy = 0.005$ m. All upscaling simulations ran for 228 days (114 timesteps) and all upsampling simulations ran for 86 days (43 time steps). It should be noted that for the gaining and losing simulations the in-situ hyporheic flux data (Fig. 1) only extended to 170 days. We therefore applied constant flow boundary conditions to the last 58 days. The upscaling simulations were also different from the upsampling simulations in that they are based on heterogeneous permeability distributions whereas the upsampling simulations contain homogeneous permeability distributions. Sample permeability fields for the upscaling simulations are given in Figure 3.

2.1.3. Simulation Variables

To train and test our model on a large variety of simulations, we added random variations to all variables of the simulations. The primary simulation variables, as well as their voxel-specific min, max, and mean values across all simulations, are given in Table 1. The average spatial distributions (in time and across all simulations) of the output features of the 1x2 meter

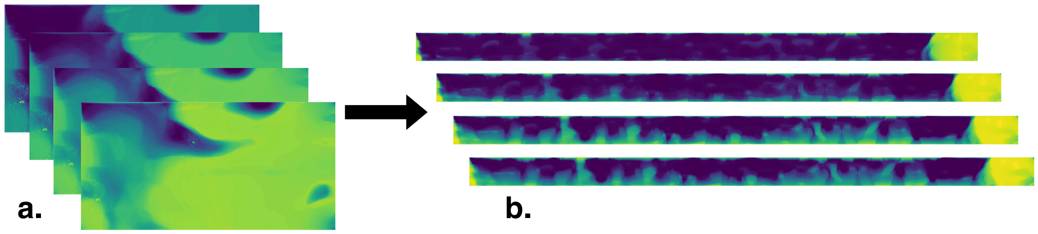

Figure 2: Sample ground truth snapshots of the 1x simulation (a) and 10x simulation (b) outputs. From top to bottom, the snapshots represent normalized biomass concentrations at $t = 40, 80, 120$, and 160 days. The primary motivation of this work is to provide a model that allows accurate mapping from 1x to 10x.

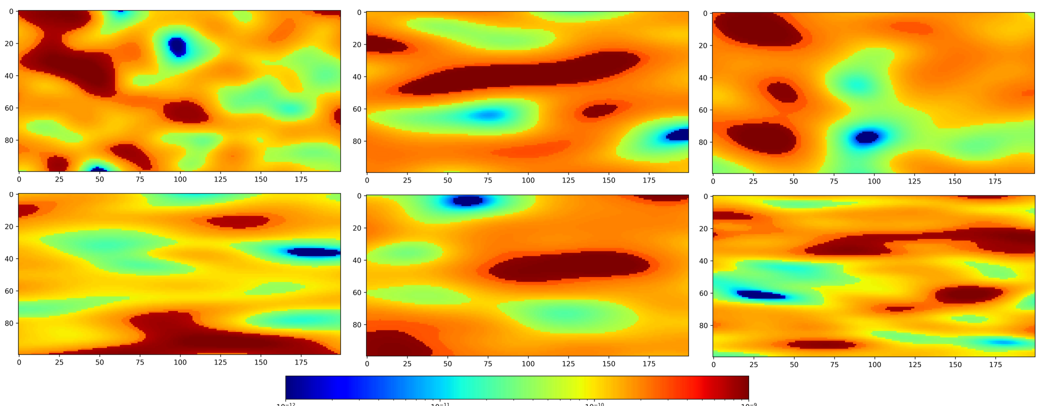

Figure 3: Sample heterogeneous permeability distributions (at t=0) used in simulations featured in this study. Heterogeneous permeability fields were used in the simulations generated to train and test our upscaling models, but for the case of our upsampling models, homogeneous permeability fields were used.

Table 1: Description of variables and their ranges of possible voxel-specific values used in the simulations. From left to right, these variables represent biomass, electron donor (molasses), and chromium concentrations, horizontal velocity, vertical velocity, pressure, temperature, porosity, permeability, biomass crowding parameter, and biomass growth parameter.

Var	B	ED	Cr(VI)	V_y	V_x	P	T	ϕ	k	α	λ_b
Units	$\frac{mol}{m^3}$	$\frac{mol}{L}$	$\frac{mol}{L}$	$\frac{m}{hr}$	$\frac{m}{hr}$	Pa	°C	-	m^2	-	-
Min	1e-10	1e-20	1e-20	-632	-486	-1214	4.8	1e-4	1e-15	0.5	1e-5
Max	765	5.5e-3	7.6e-3	671	651	7099	24.9	0.6	1.1e-9	3.0	1e-4
Mean	58	8.1e-6	1.4e-5	-5.8-2	-1.4e-2	786	11.5	0.13	2e-10	2.8	1e-5

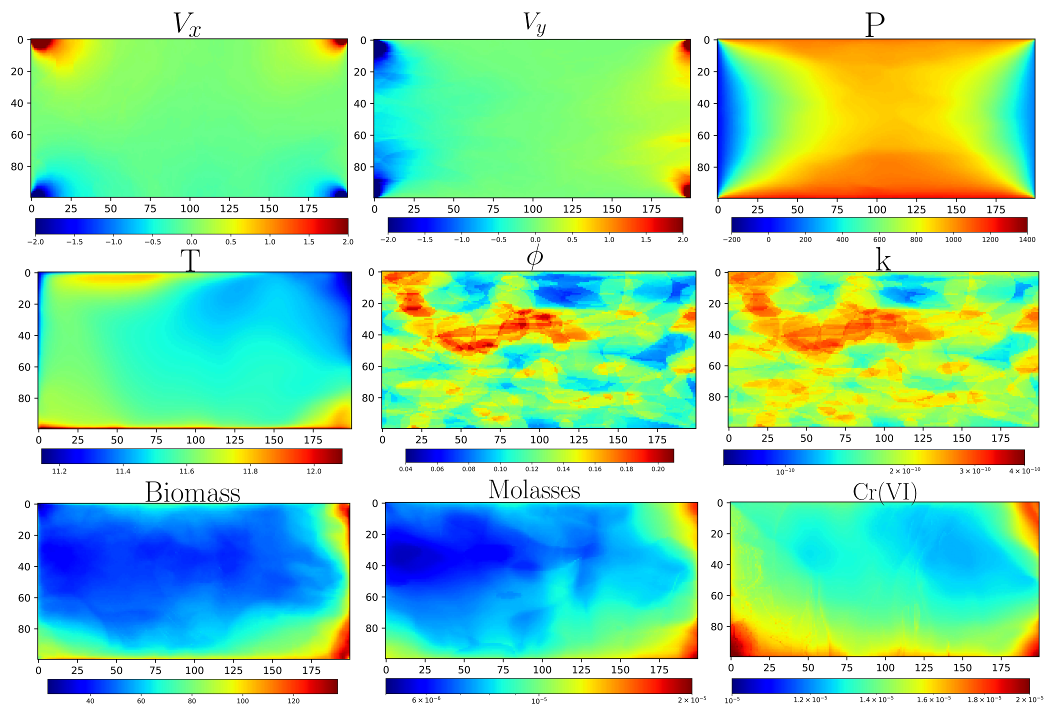

Figure 4: Mean spatial distributions for all simulations for (going from left to right and top to bottom) V_x (Darcy velocity in the vertical direction, measure in meters/hr), V_y (Darcy velocity in the direction of flow parallel to the river), pressure (Pa), temperature (°C), porosity, permeability (m^2), biomass (mol/m^3), molasses (the electron donor, measured in mol/L), and Cr(VI) (mol/L).

upscaling simulations are shown in Figure 4. Molasses, biomass, and Cr(VI) all have similar distributions due to their coupling via chemical equilibria. In addition to the features listed in Table 1 and Figure 4, less consequential features that varied between simulations included S_c, S_d1, and λ_c, which can all be classified as biomass growth parameters.

2.2. Deep-Learning-Based Upscaling

2.2.1. Model Architectures

Our initial model selection process was to look for published architectures that have been shown to be effective for spatiotemporal data [26, 27, 28]. However, our input and output tensors have shape [b, t, h, w], where b is batch size, t is the temporal dimension, and h and w are spatial dimensions. Given the irregular shape of our inputs ([b, 114, 100, 200]), and the large shape of our outputs ([b, 114, 100, 2000]) for the upscaling task, we found that the available spatiotemporal models, which are often used for classification or object detection/tracking in video data, either were too large, or would not work well with our input shape. Thus, we moved to a smaller and simpler MLP-based structure of our own design (Fig. 5a). This architecture takes in a 4D input (including batch size) and passes it through a series of linear layers with nonlinear activation functions to progressively increase the size of the final dimension to the desired number. The best number and sizes of linear layers, and the best activation function, were determined via automated hyperparameter tuning with Optuna [29]. After optimizing the structure of the MLP, we used ablation experiments with different variations of the first layer to determine the best method for initial upscaling.

For both the upscaling and upsampling models, we also investigate the impact of attention on model performance. Specifically, we integrate efficient channel attention (ECA) [30], and a novel attention method (STAM), into the optimized MLP structure after the first layer (Figs. 5b & 5c). Our efficient channel attention method uses 1D convolution in the temporal dimension, allowing the model to focus on more relevant temporal features. STAM uses convolutions in multiple dimensions (fully described in section 2.2.2) to improve focus on task-relevant spatial and temporal features. The resulting architecture with the inclusion of STAM is called STAMNet. For the rest of the paper, we refer to the upscaling version of STAMNet as STAMNet-Upscale, and the upsampling version of STAMNet as STAMNet-Upsample. STAMNet-Upsample has a different architecture than STAMNet-Upscale because the task of upsampling requires a doubling in size for both of the spatial dimensions (Fig. 5c). At a basic level, STAMNet-Upscale increases the last spatial dimension by 10x, whereas STAMNet-Upsample increases both spatial dimensions by 2x.

2.2.2. Spatiotemporal Attention Module (STAM)

The STAM architecture (Fig 6) applies attention across multiple dimensions of the input tensor. It consists of four main branches (M1 - M4 from top to bottom) that process the input in different permutations, allowing the network to capture dependencies across various dimensions.

The four attention branches can be summarized as follows - M1 processes the input along the temporal dimension through convolutional layers and reduces the size of the width dimension through a linear layer, which results in an attention map of shape [b, t, h, 1]. M2 processes the input along

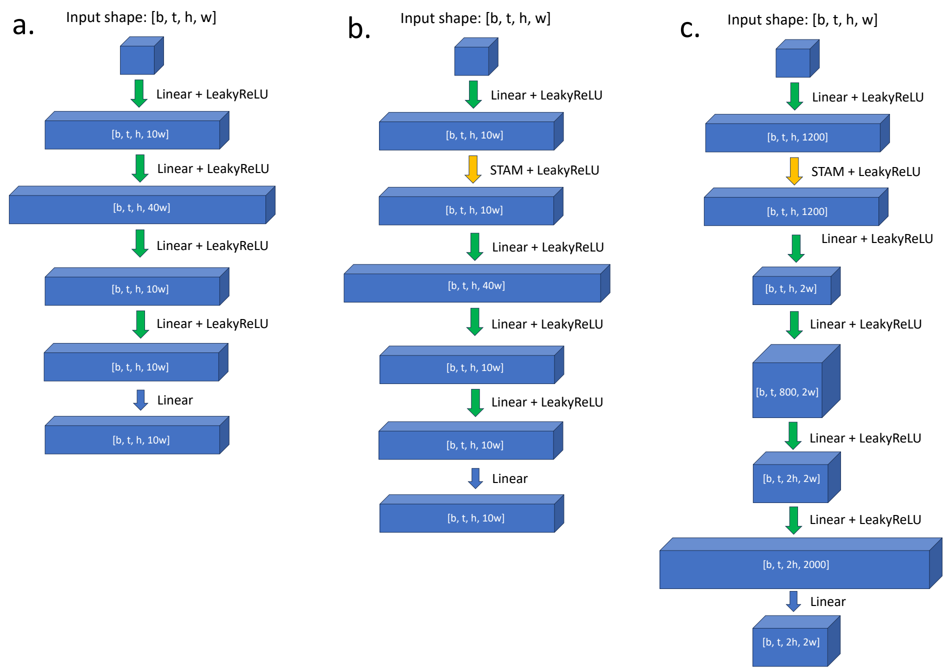

Figure 5: Model architectures for the optimal MLP (a), STAMNet-Upscale (b), and STAMNet-Upsample (c). Each block represents an intermediate output stage, and the arrows represent the layers of the model. The optimal MLP (a) and STAMNet-Upscale (b) take inputs of shape [b, t, h, w] and return outputs of shape [b, t, h, 10w]. STAMNet-Upsample takes inputs of shape [b, t, h, w] and returns outputs of shape [b, t, 2h, 2w]. STAM is a modular attention method that returns an output with the same shape as the input. The architecture of STAM is given in Figure 6. For STAMNet-Upsample (c), permutations are used after the 2nd and 4th linear layers to have the appropriate dimensions. This model takes an input of shape [b, t, h, w] and gives an output of shape [b, t, 2h, 2w].

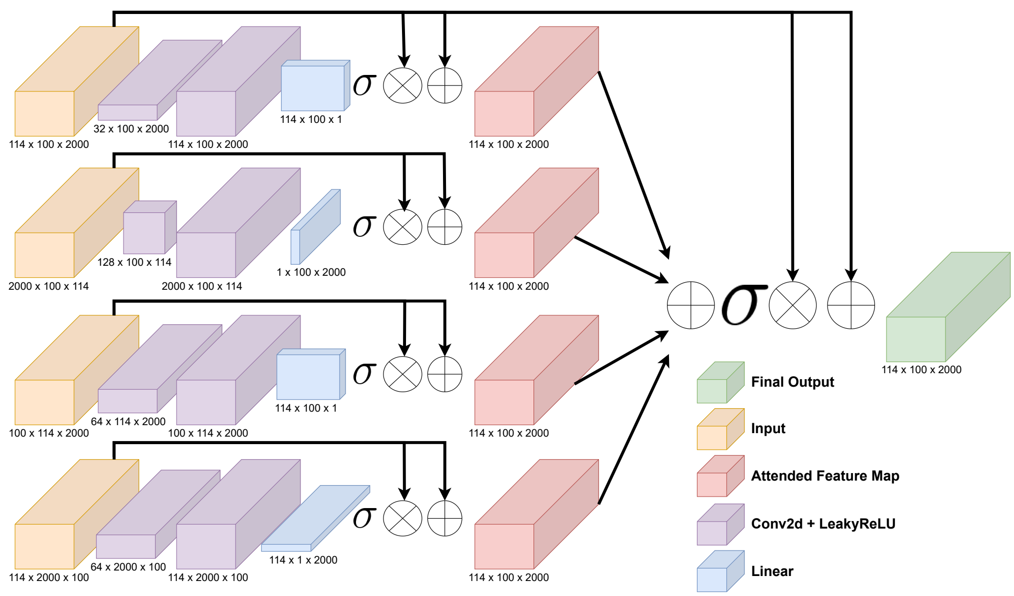

Figure 6: Architecture of the Spatiotemporal Attention Module (STAM). The model consists of four separate attention arms (M1-M4 from top to bottom). Each attention arm has two 2D convolutional layers and one linear layer followed by sigmoid activation, then multiplication and addition with the original input. The attention arms differ in their shapes, which results in feature maps that are able to capture complex cross-dimensional relationships. The attended feature maps from each attention arm are then averaged, passed through a sigmoid layer, multiplied by the input and finally added to the input to get the final output of the modular attention method.

the width dimension through convolutional layers and reduces the size of the time dimension (or the horizontal length being upscaled), which results in an attention map of shape [b, 1, h, w]. M3 processes the input along the height dimension and results in an attention map of shape [b, t, h, 1], and M4 processes the input along the temporal dimension and results in an attention map of shape [b, t, 1, w]. M3 has a similar structure to M1 (and the same output shape) except it processes the height dimension through convolution instead of the weight dimension. Each branch follows a similar pattern: Conv2D (5x5) → LeakyReLU → Conv2D (1x1) → LeakyReLU → Linear → Sigmoid. The output of each branch is multiplied with the input and the resulting product is added back to the input, creating two levels of residual connections. The outputs from all branches are then averaged and passed through a sigmoid activation function and multiplied and then added to the input to get the final attention map, thus creating additional residual connections.

STAM incorporates several architectural features that enhance its ability to map spatiotemporal relationships. By processing the input tensor along different dimensions, it captures complex spatial-temporal dependencies that simpler attention mechanisms or non-attentive models might overlook. The combination of 5x5 and 1x1 convolutions enables STAM to integrate both local and global context within each dimension [31, 32, 33]. Through residual connections and a final aggregation step, the model adaptively refines features, highlighting important patterns while attenuating less relevant information [34, 35]. The incorporation of LeakyReLU activations and dropout (in M2) introduces non-linearity and regularization, potentially enhancing the model's generalization capabilities [36]. Furthermore, by processing the input through different permutations, STAM generates complementary attention maps, effectively capturing diverse data patterns [37]. Lastly, the addition of the input to the attention-weighted features preserves original information while facilitating the learning of cross-dimensional representations, thus providing a comprehensive approach to spatiotemporal feature extraction and refinement.

2.2.3. Training, Validation and Testing Process

Although the RT simulations contain multiple output features (Fig. 4), we chose to focus on upscaling biomass, Cr(VI), and molasses. All other variables either have little variation between 1x and 10x scale simulations (such as temperature and pressure) or can be easily upscaled through physics-based methods (such as flow [38, 39, 40] and permeability [41]. With about 48 GB of VRAM, models could be developed to upscale all three variable at once. However, we were restricted to 24 GB of VRAM, and at this amount of VRAM we weren't able to effectively train multi-feature models. Thus, we trained a suite of models that separately upscale our three target variables.

For the upscaling task, 48 pairs of simulations were used for training, 8 pairs of simulations were used for validation, and 13 pairs of simulations were used for testing. Validation scores were used to optimize hyperparameters and determine layer placement within STAMNet. Once the best model architectures were determined through hyperparameter optimization and ablation studies, the validation data was also used for training, resulting in 56 pairs of training simulations and 13 pairs of testing simulations for the final calculation of scores. We used the AdamW optimizer, dropout of 0.2 after the first linear layer (or after the attention layer for models with attention), and a learning rate (lr) between 1.6e-4 and 4.6e-4 with a cosine annealing warm restarts scheduler. For biomass upscaling we used a lr of 1.6e-4 and trained for 25 epochs, for biomass upsampling we used a lr of 4.6e-4 and trained for 900 epochs, and for ED and CR(VI) upscaling we used a lr of 3.0e-4 and trained for 110 epochs. The number of epochs used for each feature was determined based on when the validation set stopped showing improvement. For all biomass upscaling experiments, we trained and tested each model type 14 times and report the averages of each performance metric (MSE, MAE, and R^2). Each metric is calculated between all elements of the output tensor (x) and the ground truth (y). For example, the MAE is the sum of errors between each element of x and y divided by the number of elements in x and y. We also plot the mean time series and spatial distributions (over all simulations) for each model to provide a visual understanding of the prediction errors. Specifically we use time series to investigate the average temporal distributions of the predictions of biomass, molasses (ED), and chromium for a simple interpolation model, an MLP+ECA model, and our STAMNet model. The spatial distributions are presented in two ways. The blue-green-yellow spatial distributions show the absolute concentrations of the feature in question, while the blue-white-red spatial distributions show the difference between the ground truth and the prediction for that particular model. For this visual analysis, we use a simple ensemble of the best-scoring variations of each model.

For the upsampling task, 40 pairs were used for training, 8 for validation, and 12 for testing. After optimal model structures were determined, the 8 simulation pairs used for validation were included in the training set, resulting in a final 48 simulation pairs for training and 12 simulation pairs for testing. All results for the upsampling task are a comparison of the average of 8 separately trained and tested models. For both upscaling and upsampling, the loss function used for training was $MSE + 0.6 \times MAE$, which was used over a standard MSE loss function as we found that only using MSE tends to result in a higher degree of overfitting. Furthermore, we found Huber loss to not weight the MAE strongly enough, which resulted in decreased MAE and R^2 scores.

3. Results and Discussion

3.1. Ablation Experiments

To determine the best simple method of upscaling, we experimented with three model variations (Table 2). The structure of the linear model is shown in Figure 5a. This structure was determined through optimization of validation scores via Optuna. In the linear model, the first layer is a linear layer that increases the size of the final dimension of the input by 10x. To reduce the number of parameters, or have roughly the same number of parameters with a deeper first layer, we tried to replace the first linear layer with a 10x interpolation layer and a 20x interpolation layer. Since simple interpolation often allows for reasonably accurate upscaling, the interpolation scheme implemented here is expected to be a parameter-efficient way to upscale the final dimension. The 10x interpolation layer takes the input of shape [b, t, h, w] and outputs a tensor of shape [b, t, h, 10w], while the 20x interpolation layer takes the same shape of input and outputs a tensor of shape [b, t, h, 20w]. Thus, the interpolation 10x model has the same structure as the linear model besides the 1st layer, which is instead a 10x repeat interleave layer. Similarly, the 20x interpolation model has an initial layer that interpolates the final dimension of the input to 20x size. Because the linear layer of the optimal MLP upscales the final dimension to 10x, the 20x model has a slightly different structure of second layer as it takes an input of [b, t, h, 20w]. The linear model performed best in the MAE and R^2 metrics. Thus, although interpolation allows for model parameter savings, it is not much, and the reduced accuracy is not worth these savings in most cases, so we developed STAMNet on top of this optimal linear architecture.

Table 2: Ablation experiments for biomass upscaling to determine the best method of increasing dimension size. MAE is given in $\frac{mol}{m^3}$ and MSE is given in $\frac{mol}{m^3}^2$.

	Linear	Interpolation (10x)	Interpolation (20x)
MSE	2508	2471	2471
MAE	21.62	22.25	22.02
R^2	0.897	0.880	0.875

3.2. STAMNet-Upscale Performance

3.2.1. Biomass Upscaling

The results of our upscaling models for the biomass prediction task are given in Table 3. The interpolation model here is different than the interpolation models used in the ablation experiments. In the ablation experiments, the interpolation was used as an initial layer of a model with multiple linear layers and an activation function after the interpolation. For the interpolation model in table 3, there are no linear layers after the interpolation. In other words, it is just a simple interpolation of the final dimension, which is the most simple and rapid way to generate reasonably accurate results for the task of upscaling as defined in this paper. This interpolation model can also be though of as a simple reproduction of the baseline scale (but with extended dimensions), meaning that the errors for the interpolation model are a proxy for the differences between 1x and 10x scale. The simple MLP is a one-layer MLP that increases the size of the final dimension by 10x. The optimal MLP is the fully optimized MLP structure given in Figure 5a. The structure of STAMNet, our proposed best-performing model, can be seen in Figure 5c. The MLP+ECA model has the same structure as STAMNet, but with the ECA attention module instead of the STAM attention module.

For all models tested, STAMNet-Upscale shows the strongest performance by a statistically significant margin for the both MSE and R^2 metrics. Both models with attention modules outperform the optimal MLP, further indicating that attention is a useful tool for developing robust upscaling model architectures. STAMNet-Upscale performs better than the MLP+ECA model, indicating that cross-dimensional feature refinement offers performance benefits over single dimensional (temporal) feature refinement. All trained models perform better than simple interpolation, showing the general benefit to the approach of using deep learning for upscaling of reactive transport simulations.

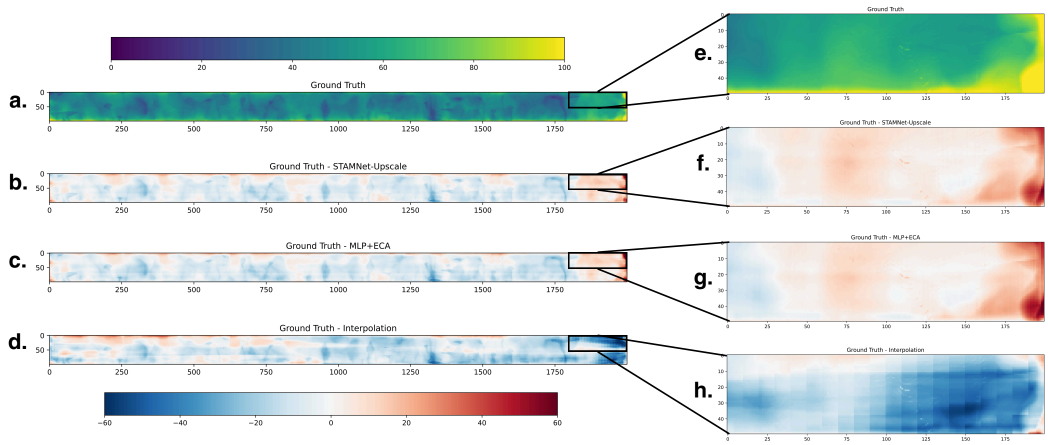

Figure 7: Spatial error distribution of biomass averaged over all test simulations and time steps. (a) Ground truth spatial distribution. (b) Ground truth minus output from STAMNet-Upscale. (c) Ground truth minus the MLP+ECA model. (d) Ground truth minus simple interpolation. (e-h) Zoomed in versions of a-d. These figures show that STAMNet has difficulty predicting fine spatial variations but is more accurate than a simple interpolation.

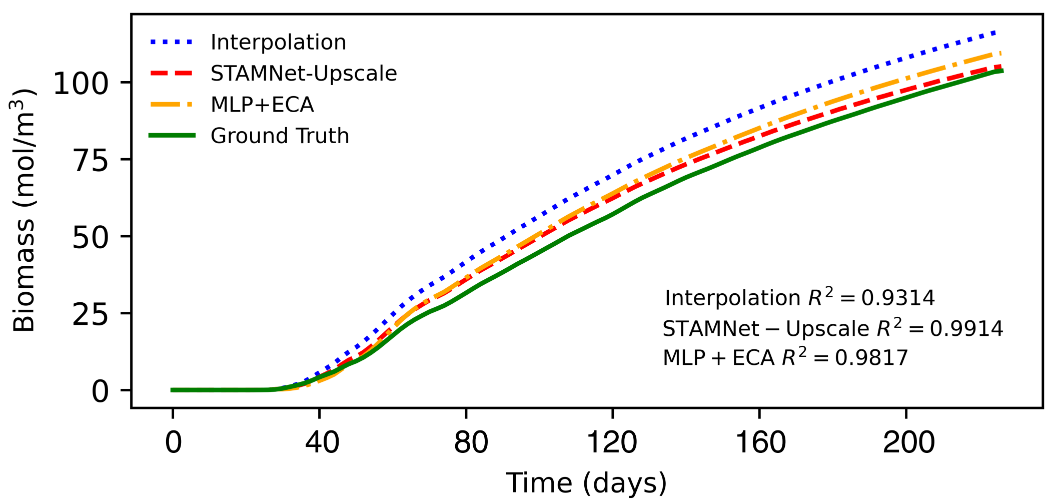

Figure 8: Biomass time series averaged over all upscaling test simulations and time steps. The blue dotted line corresponds to the time series for a simple interpolation of the input, the red dashed line corresponds to the output for STAMNet-Upscale, the yellow dash-dotted line corresponds to the output for the MLP+ECA model, and the green dash-dotted line corresponds to the ground truth (i.e., the time series of the 10x scale simulation output). This figure shows STAMNet outperforms simple interpolation and achieves a high level of accuracy in terms of time series prediction.

To further investigate the performance of our different models for the task of upscaling biomass, we plot the spatial error (Fig. 7) and the mean time series error (Fig. 8). The spatial errors show that the simple interpolation (Fig. 7d), the MLP+ECA (Fig. 7c), and STAMNet-Upscale (Fig. 7b) all fail to capture fine variations in the ground truth spatial distribution. Instead, they achieve a low MSE/MAE by averaging out the variabilities in space. This is to be expected, however. Without a method that specifically constrains the spatial distributions of biomass concentrations, the model lacks the ability to predict exactly what the upscaled version will look like, so the model just makes an average guess. In other words, the neural nets may learn to approximate the vertical variability in biomass well, since this doesn't change much between small and large-scale simulations, but have no ability to predict the horizontal variability in biomass as this may change significantly based on scale and more strongly depends on the differences between the small and large-scale permeability fields. To compensate for this lack of knowledge, the neural nets make predictions that represent averages across many horizontal voxels. We experimented with loss functions to try to add this constraint to the spatial distribution of the outputs, but found it had too negative of an impact on the outputs of time series distributions and did not improve the accuracy of the spatial distributions (either in exact value or "look") enough to warrant further investigation. Thus, although there are some differences in the spatial error between different models, no model we tested provides an adequate representation of physically realistic spatial variations, and more robust techniques are needed to achieve high-fidelity spatial predictions.

In terms of comparison between the methods, STAMNet-Upscale and the MLP+ECA clearly outperform the interpolation, which can be better seen from the zoomed-in sections of the spatial error distributions (Figs. 7e-7h). There is a very slight difference between error for STAMNet-Upscale and the MLP+ECA, but it is essentially negligible with regards to the overall accuracy of the predictions of spatial distributions. One big difference between the interpolation and the attention-based neural nets is that the interpolation model greatly overcalculates biomass concentrations, especially near the right boundary of the domain. The right boundary of the domain is often a source of nutrients and thus a location of dense biomass growth. In the 1x2 meter simulation, these nutrients are able to reach into and cause biomass growth in about half of the domain, meaning a simple interpolation to the 1x20 meter simulation leads to high biomass concentrations that extend too far into the domain. The primary reason for this significant error in the simple interpolation model is that the 1x and 10x scale simulations have the same pressure boundary conditions, meaning they have different pressure gradients. Specifically, the 1x simulation has a pressure gradient 10 times greater than the 10x solution, meaning horizontal flow results in less biomass growth at 10x scale in terms of the percentage of the domain. The neural nets are able to provide more accurate mapping between the two scales, although they still show significant spatial errors at the right-hand boundary due to their tendency to average local variations in concentration.

In addition to our analysis of the spatial errors of the simple interpolation method and attention-based models, we also investigate their performance in terms of the average time series prediction (Fig. 8). Unlike the spatial distributions, all models perform quite well at the task of capturing the average upscaled time series. Both attention-based models clearly outperform a simple interpolation, and STAMNet-Upscale slightly outperforms the MLP+ECA model.

3.2.2. Low and High-Concentration Biomass Upscaling

To further refine our general investigation of the upscaling potential of STAMNet for RT simulations, we split this analysis up to investigate performance on high-concentration and low-concentration simulations. Of the 13 test simulations, 5 simulations can be categorized as high-concentration (mean biomass greater than 50 mol/m^3), and 5 simulations can be categorized as low-concentration (mean biomass less than 15 mol/m^3). The spatial errors for both the low and high concentrations (Fig. 9) show the same trends as the spatial errors for the full set of results (Fig. 7). For the low concentrations, STAMNet-Upscale (Fig. 9b) and the MLP+ECA model (Fig. 9c) are completely indistinguishable, and both clearly outperform the interpolation (Fig, 9d). For the high concnetrations, there is similarly very little difference between the spatial errors for STAMNet (Fig. 9f) and the MLP+ECA model (Fig. 9g). Both attention-based neural nets outperform simple interpolation (Fig. 9h), although similar to the full set of results (Fig. 7) for biomass upscaling, these differences are negligible compared to the overall error of the spatial distributions.

The time series plots for the low and high-concentration upscaling (Fig. 10) reveal slightly more interesting deviations from the analysis of all test simulations. The low-concentration time series (Fig. 10a) shows dramatically better performance for the attention-based neural nets when compared to the simple interpolation. The high-concentration time series (Fig. 10b), on the

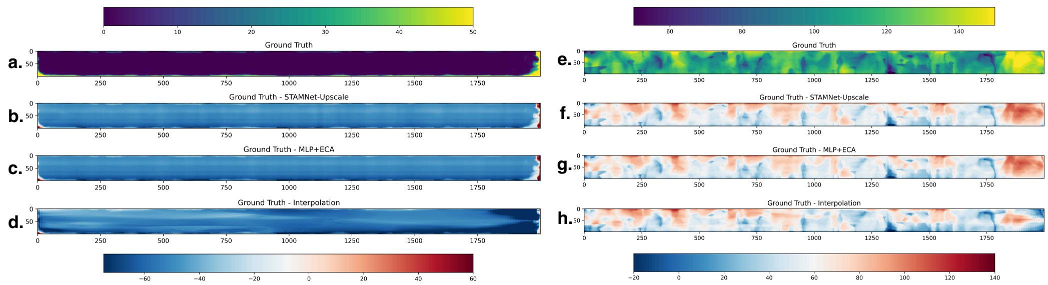

Figure 9: Spatial error distribution of biomass averaged over low and high-concentration test simulations.(a) Ground truth spatial distribution for low-concentration simulations.(b) Ground truth minus output from STAMNet-Upscale for low-concentration simulations.(c) Ground truth minus the MLP+ECA model for the low-concentration simulations. (d) Ground truth minus simple interpolation for low-concentration simulations. (e-h) High-concentration versions of a-d. These figures show greater difference between the spatial errors of the interpolation and our trained networks for low-concentration simulations than high-concentration simulations.

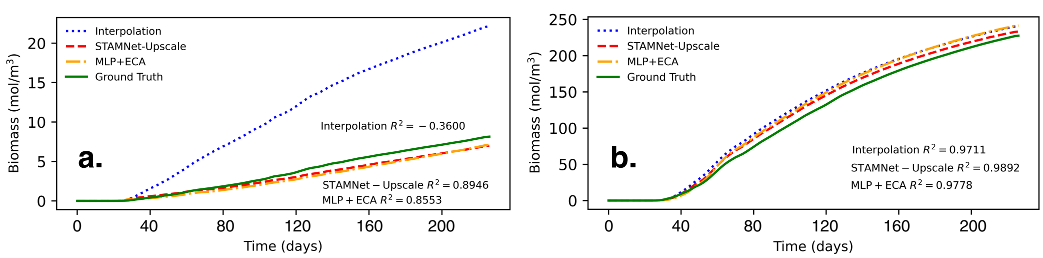

Figure 10: Biomass time series averaged over low (a) and high-concentration (b) upscaling test simulations. This figure shows STAMNet-Upscale outperforms simple interpolation and our MLP+ECA architecture, especially for low-concentration simulations.

Table 3: Results for biomass upscaling. Each model was trained and tested 14 times, and values here correspond to the average scores over all 14 model iterations. MAE of biomass is given in $\frac{mol}{m^3}$ and MSE is given in $\frac{mol}{m^3}^2$. Values in bold indicate statistically significantly better performance than all other models.

	Interpolation	Simple MLP	Optimal MLP	MLP + ECA	STAMNet-Upscale
MSE	4089	2691	2508	2525	**2480**
MAE	26.42	28.08	21.62	**21.35**	21.60
R^2	0.727	0.903	0.897	0.892	**0.925**

other hand, shows relatively small differences between each model. Thus, our results indicate that a simple interpolation is a generally accurate way to upscale high-concentration RT simulations of biomass growth, but not for low-concentration simulations as it results in a large amount of overprediction of biomass concentrations. As previously discussed, this overprediction is due to the differences in pressure gradients at low and high scales, which appear to be especially relevant in low-concentration simulations.

Across all biomass upscaling experiments, our results show that compared to models without attention, STAM can selectively focus on important features in the input, potentially leading to better performance on tasks that require understanding of complex spatial-temporal relationships. Compared to ECA, which focuses on attention in the temporal dimension, STAM provides a more comprehensive attention mechanism that considers both spatial and temporal dimensions. STAM generally performs better than ECA, which indicates the cross-dimensional attention, which improves feature mapping in both the spatial and temporal dimensions, is advantageous for the task of upscaling.

3.2.3. Molasses and Cr(VI) Upscaling

We further show the strong performance of STAMNet by upscaling RT simulations for the molasses (electron donor - ED) and Cr(VI) features. We find that STAMNet significantly outperforms the simple interpolation and MLP+ECA model for ED prediction in the MSE and R^2 metrics (Table 4). Furthermore, the MLP+ECA model also significantly outperforms the simple interpolation in all metrics, which once again shows the benefit of neural architectures, and especially those with attention, for the task of upscaling. Specifically, these results show that the benefits of our neural architectures for upscaling are not restricted to biomass, and can be extended to other features. The results for Cr(VI) similarly show high upscaling per-

Table 4: Results for upscaling experiments with molasses (ED) and Chromium. MAE is given in $\frac{mol}{L}$ and MSE is given in $\frac{mol}{L}^2$. Values in bold indicate statistically significantly better performance than all other models.

ED	Interpolation	MLP + ECA	STAMNet-Upscale
MSE	1.20×10^{-4}	2.77×10^{-5}	$\mathbf{2.62 \times 10^{-5}}$
MAE	2.72×10^{-3}	1.81×10^{-3}	1.80×10^{-3}
R^2	0.2937	0.888	**0.913**
Cr(VI)	Interpolation	MLP + ECA	STAMNet-Upscale
MSE	5.32×10^{-4}	1.02×10^{-4}	8.92×10^{-5}
MAE	3.31×10^{-3}	2.45×10^{-3}	2.23×10^{-3}
R^2	0.9093	0.9508	0.9227

formance for both STAMNet-Upscale and the MLP+ECA model. Also, in this case, the MLP+ECA model outperforms STAMNet-Upscale in the R^2 metric, but performs worse than STAMNet-Upscale for the MSE and MAE metrics (although none of these differences are significant), indicating that the differences in performance between STAM and ECA may depend on the particular task. Thus, in addition to providing trained models for biomass, ED, and Cr(VI) upscaling, we provide multiple frameworks with which future researchers can train their own upscaling models for specific features they may be interested in. Although we find STAMNet to outperform our optimized MLP+ECA model, we encourage researchers to run their own experiments with their data to determine which model works best for their task.

Similar to our results for the biomass upscaling, we also present spatial and time series errors for chromium and molasses (Fig. 11). Like all the spatial errors for biomass, we find that STAMNet-Upscale and the MLP+ECA model outperform the simple interpolation but are not able to capture the fine-grained details of the spatial distributions for molasses (Fig. 11c) and chromium (Fig. 11d). For the mean time series comparison, we see STAMNet-Upscale and the MLP+ECA model perform equally well at molasses upscaling (Fig. 11a). Surprisingly, although the MLP+ECA model gives a higher R^2 for Cr(VI) when the metric is calculated as the mean over the set of R^2 values for each time series, when the R^2 is calculated from the mean time series (i.e. the time series is averaged over all siulations then R^2 is calculated), we see STAMNet-Upscale has a slightly higher R^2 (Fig. 11b). This indicates that MLP+ECA is more accurate for the Cr(VI) upscaling

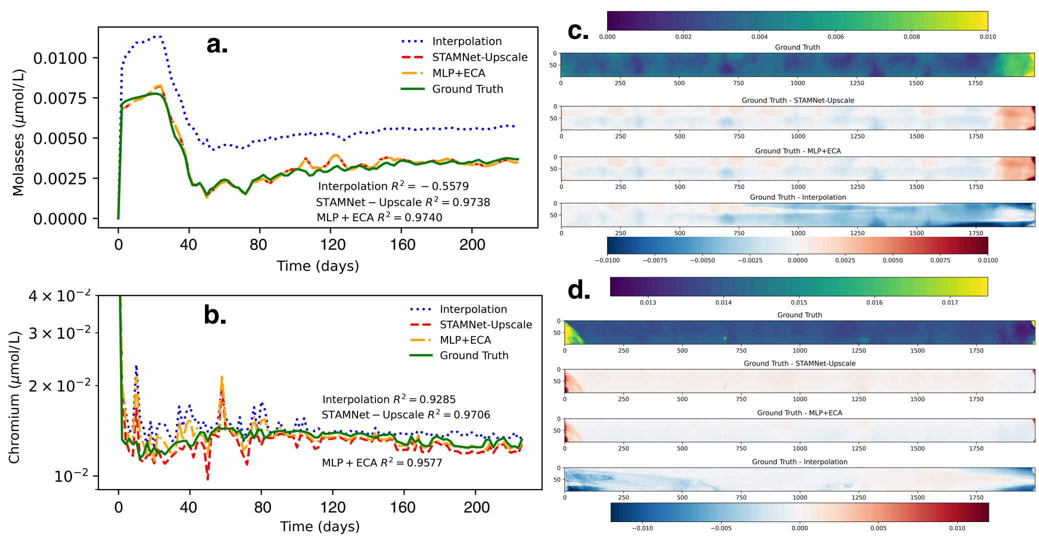

Figure 11: STAMNet-Upscale, MLP+ECA, and interpolation performance for (a) molasses time series, (b) chromium time series, (c) molasses spatial error distributions, and (d) chromium spatial error distributions. STAMNet outperforms interpolation in the molasses time series and spatial distribution and the chromium spatial distribution, but it is nearly indistinguishable from the interpolation for the chromium time series.

of any given simulation, but that STAMNet-Upscale will be more accurate when considering the mean value of a variety of simulations. The simple interpolation performs well for Cr(VI), but not for molasses, further showing its inconsistent performance compared to that of the attention-based neural nets.

3.3. STAMNet-Upsample Performance

In addition to our investigation of upscaling in the sense of increasing the lateral domain of the simulation output, we also use another variant of the STAMNet architecture (Fig. 5c) to increase the resolution of the simulation

Table 5: Results for upsampling experiments. MAE is given in $\frac{mol}{m^3}$ and MSE is given in $\frac{mol}{m^3}^2$.

	Interpolation	Simple MLP	Optimal MLP	MLP+ECA	STAMNet-Upsample
MSE	73.50	54.45	28.58	29.76	**26.22**
MAE	1.443	1.557	1.263	1.268	**1.238**
R^2	**0.9997**	0.9987	0.9984	0.9987	0.9986

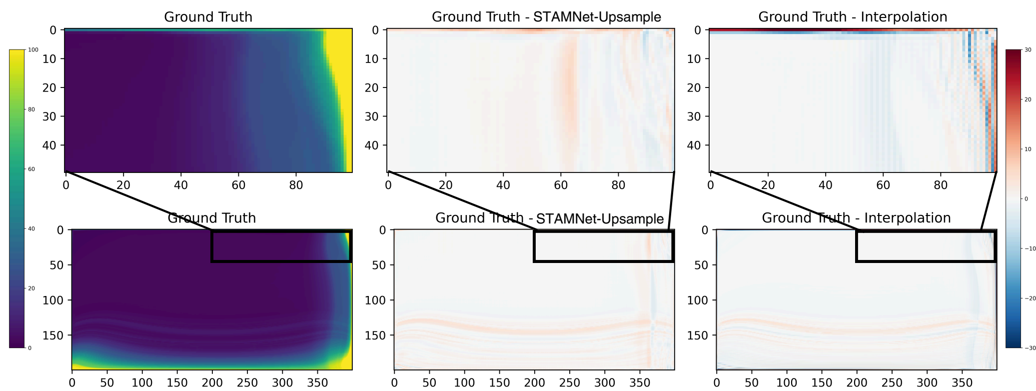

Figure 12: Spatial error distribution of biomass averaged over all test simulations and time steps for the upsampling task. (a) Ground truth spatial distribution. (b) Ground truth minus output from STAMNet-Upscale. (c) Ground truth minus simple interpolation. (d-f) Zoomed in versions of a-c. These figures show the superior spatial performance of STAMNet over interpolation for the task of upsampling, especially near the boundaries of the domain.

output, which we refer to as the task of upsampling. We find that STAMNet-Upsample shows significantly better performance than all other models for the upsampling task in MSE and MAE, although a simple interpolation gives the best performance for R^2 (Table 5). Considering we trained and tested on a smaller number of simulations for the task of upsampling compared to the task of upsampling, and that we only used homogeneous permeability fields for the upsampling simulations, it is likely that the upsampling task was not difficult enough to result in large performance differences between models. Furthermore, some degree of overfitting on the MSE and MAE during training may have resulted in a model that focuses more on the spatial aspects of the upsampling task than the average temporal ones. Looking at the spatial distributions for STAMNet-Upsample and the interpolation (Fig. 12), we can see that STAMNet-Upsample does indeed have more accurate spatial approximations (especially at the domain boundaries). Thus, while it could be argued the simple interpolation may be more appropriate for tasks that don't care about the accuracy of the spatial distribution, for tasks where spatial accuracy is important, STAMNet-Upsample is clearly advantageous to simple interpolation. We also find that the addition of STAM to the optimized MLP improves performance, but the addition of ECA to the optimized MLP generally decreases performance. These results contrast those of the upscaling task, where both ECA and STAM were found to improve MLP performance. As seen by the high R^2 of all models for the average time series, the task of upsampling does not result in large differences in the average temporal trends. Thus, ECA, which uses 1D convolution to improve temporal feature extraction, results in a tiny improvement in R^2 (relative to the optimal MLP), but causes a decrease in MAE and MSE due to the extra focus on temporal features. This trend is not observed in the upscaling results (Table 3), however, as ECA shows significant improvements to MAE for biomass upscaling. Given the significant differences in task and input tensors, these differences are likely a result of differences in the task-specific performances of each architecture and imperfect training hyperparameters. Thus, although ECA generally doesn't perform as well as STAMNet-Upsample for the task of upsampling, it is possible these results would be different if hyperparameters were individually tuned for each model, which strengthens our suggestion for future researchers to experiment with both the STAMNet and MLP+ECA architectures.

4. Conclusions

This study presents STAMNet, a novel deep learning architecture for upscaling and upsampling reactive transport simulations in the hyporheic zone with specific applications to Monte-Carlo-style investigations of bioremediation. Our results demonstrate that STAMNet outperforms traditional interpolation methods and simpler neural network architectures across multiple tasks and metrics. STAMNet-Upscale significantly improved upon simple interpolation and optimized MLP models for biomass upscaling, achieving the highest R^2 (0.925) and lowest MSE among tested models. Furthermore, the spatiotemporal attention module (STAM) consistently outperformed efficient channel attention (ECA), indicating the benefits of cross-dimensional feature refinement for reactive transport modeling. Notably, STAMNet showed robust performance across different concentration regimes, with particularly strong improvements over interpolation for low-concentration simulations. The architecture performed well with other reactive transport variables, showing significant improvements for molasses (electron donor) upscaling and competitive performance for chromium upscaling. Furthermore, STAMNet-Upsample demonstrated superior performance in increasing simulation resolution, particularly in capturing spatial details more accurately than simple interpolation. By enabling rapid upscaling and upsampling of simulations, this approach has the potential to accelerate research in hyporheic zone processes and enhance our ability to quickly design bioremediation approaches at large scales. As the field continues to evolve, the integration of advanced deep learning architectures like STAMNet with domain-specific knowledge promises to unlock new possibilities in environmental modeling and decision-making.

While STAMNet shows promise for accelerating and enhancing reactive transport simulations, some limitations should be noted. The current implementation struggles to capture fine-grained spatial variations in upscaled simulations, instead producing averaged distributions. Additionally, the model's performance may vary depending on the specific reactive transport variable being predicted, as seen in the differences between biomass, molasses, and chromium results. It's also important to acknowledge that the study focused on a specific bioremediation scenario, and further testing is needed to confirm generalizability to other subsurface environments and reactive transport systems.

Future research directions should address these limitations and expand upon the current work. Investigators should explore methods to incorporate physical constraints or multi-scale approaches to improve the spatial fidelity of upscaled predictions. Extending the model to handle a wider range of reactive transport variables and scenarios, including more complex biogeochemical reactions, constant pressure gradients, and heterogeneous subsurface environments, would further enhance its applicability. The integration of STAMNet with physics-based models to create hybrid approaches that leverage both data-driven and mechanistic insights is also an exciting avenue for development. Finally, investigating the potential of STAMNet for other spatiotemporal prediction tasks beyond reactive transport, such as climate modeling or ecosystem dynamics, could open up new applications for this innovative architecture.

CRediT authorship contribution statement

Marc Berghouse: Data curation, Formal analysis, Investigation, Software, Conceptualization, Visualization, Writing – original draft, Writing – review & editing, Methodology. Rishi Parashar: Supervision, Writing - review & editing, Project administration.

Declaration of Competing Interests

The authors declare no competing interests.

Acknowledgments

This research is based upon work supported by the U.S. Department of Energy (DOE) under award number DE-SC0019437. The second author (R.P) acknowledges career advancement support in the field of artificial intelligence based methods provided by NSF grant 2123481.

Data Availability

Data used for training and testing are available upon request. The source codes are available for downloading at the link: https://github.com/mberghouse/STAMNet

References

[1] M. B. Cardenas, Hyporheic zone hydrologic science: A historical account of its emergence and a prospectus, Water Resources Research 51 (5) (2015) 3601–3616.

[2] S. Krause, D. M. Hannah, J. H. Fleckenstein, C. M. Heppell, D. Kaeser, R. Pickup, G. Pinay, A. L. Robertson, P. J. Wood, Inter-disciplinary perspectives on processes in the hyporheic zone, Ecohydrology 4 (4) (2011) 481–499. doi:https://doi.org/10.1002/eco.176.

[3] M. Rode, M. Hartwig, D. Wagenschein, T. Kebede, D. Borchardt, The Importance of Hyporheic Zone Processes on Ecological Functioning and Solute Transport of Streams and Rivers, Springer Netherlands, Dordrecht, 2015, pp. 57–82.

[4] A. J. Boulton, T. Datry, T. Kasahara, M. Mutz, J. A. Stanford, Ecology and management of the hyporheic zone: stream–groundwater interactions of running waters and their floodplains, Journal of the North American Benthological Society 29 (1) (2010) 26–40. doi:10.1899/08-017.1. URL https://doi.org/10.1899/08-017.1

[5] S. Roy Chowdhury, J. P. Zarnetske, M. S. Phanikumar, M. A. Briggs, F. D. Day-Lewis, K. Singha, Formation criteria for hyporheic anoxic microzones: Assessing interactions of hydraulics, nutrients, and biofilms, Water Resources Research 56 (3) (2020) no–no.

[6] G. Hammond, P. Lichtner, C. Lu, R. T. Mills, Pflotran: Reactive flow & transport code for use on laptops to leadership-class supercomputers, Groundwater reactive transport models 5 (2012) 141–159.

[7] M. White, M. Oostrom, Stomp subsurface transport over multiple phases: Users guide, Tech. rep., Pacific Northwest National Lab.(PNNL), Richland, WA (United States) (1997).

[8] C. I. Steefel, S. Molins, Crunchflow, Software for modeling multicomponent reactive flow and transport. User's manual (2009) 12–91.

[9] C. I. Steefel, K. McQuarrie, Approaches to modeling of reactive transport in porous media, Reviews in mineralogy 34 (1996) 83–130.

[10] D. Su, K. U. Mayer, K. T. MacQuarrie, Parallelization of min3p-thcm: A high performance computational framework for subsurface flow and reactive transport simulation, Environmental Modelling & Software 95 (2017) 271–289.

[11] S. Yu, W. Hannah, L. Peng, J. Lin, M. A. Bhouri, R. Gupta, B. Lütjens, J. C. Will, G. Behrens, J. Busecke, et al., Climsim: A large multi-scale dataset for hybrid physics-ml climate emulation, Advances in Neural Information Processing Systems 36 (2024).

[12] O. Hennigh, S. Narasimhan, M. A. Nabian, A. Subramaniam, K. Tangsali, Z. Fang, M. Rietmann, W. Byeon, S. Choudhry, Nvidia simnet™: An ai-accelerated multi-physics simulation framework, in: International conference on computational science, Springer, 2021, pp. 447–461.

[13] Y. Sun, J. Lu, Q. Liu, W. Shuai, A. Sun, N. Zheng, Y. Han, G. Xiao, J. Xuan, M. Ni, et al., Multi-objective optimizations of solid oxide co-electrolysis with intermittent renewable power supply via multi-physics simulation and deep learning strategy, Energy conversion and management 258 (2022) 115560.

[14] R. Sun, H. Jeong, J. Zhao, Y. Gou, E. Sauret, Z. Li, Y. Gu, A physics-informed neural network framework for multi-physics coupling microfluidic problems, Computers & Fluids (2024) 106421.

[15] D. Botache, J. Decke, W. Ripken, A. Dornipati, F. Götz-Hahn, M. Ayeb, B. Sick, Enhancing multi-objective optimisation through machine learning-supported multiphysics simulation, in: Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Springer, 2024, pp. 297–312.

[16] E. Laloy, D. Jacques, Emulation of cpu-demanding reactive transport models: a comparison of gaussian processes, polynomial chaos expansion, and deep neural networks, Computational Geosciences 23 (2019) 1193–1215.

[17] E. Demirer, E. Coene, A. Iraola, A. Nardi, E. Abarca, A. Idiart, G. de Paola, N. Rodríguez-Morillas, Improving the performance of reactive transport simulations using artificial neural networks, Transport in Porous Media 149 (1) (2023) 271–297.

[18] A. M. Leal, S. Kyas, D. A. Kulik, M. O. Saar, Accelerating reactive transport modeling: on-demand machine learning algorithm for chemical equilibrium calculations, Transport in Porous Media 133 (2) (2020) 161–204.

[19] N. I. Prasianakis, R. Haller, M. Mahrous, J. Poonoosamy, W. Pfingsten, S. V. Churakov, Neural network based process coupling and parameter upscaling in reactive transport simulations, Geochimica et Cosmochimica Acta 291 (2020) 126–143.

[20] Z. Wang, I. Battiato, A deep learning upscaling framework: Reactive transport and mineral precipitation in fracture-matrix systems, Advances in Water Resources 183 (2024) 104588.

[21] J. You, K. J. Lee, Upscaling reactive transport models from pore-scale to continuum-scale using deep learning method, Geoenergy Science and Engineering 238 (2024) 212850.

[22] S. K. Hansen, S. Pandey, S. Karra, V. V. Vesselinov, Chrotran 1.0: A mathematical and computational model for in situ heavy metal remediation in heterogeneous aquifers, Geoscientific Model Development 10 (12) (2017) 4525–4538.

[23] K. Chen, X. Chen, J. C. Stegen, J. A. Villa, G. Bohrer, X. Song, K.-Y. Chang, M. Kaufman, X. Liang, Z. Guo, et al., Vertical hydrologic exchange flows control methane emissions from riverbed sediments, Environmental Science & Technology 57 (9) (2023) 4014–4026.

[24] J. Ren, B. Chen, W. Zhang, L. Men, J. Yang, Y. Li, Quantification of the temporal–spatial distributions characteristics of streambed hyporheic exchange fluxes with the seasonal variation using heat as a tracer, Environmental Earth Sciences 79 (2020) 1–19.

[25] H. Le Lay, Z. Thomas, F. Rouault, P. Pichelin, F. Moatar, Characterization of diffuse groundwater inflows into stream water (part ii: quantifying groundwater inflows by coupling fo-dts and vertical flow velocities), Water 11 (12) (2019) 2430.

[26] Z. Yang, M. Ding, B. Xu, H. Yang, J. Tang, Stam: A spatiotemporal aggregation method for graph neural network-based recommendation, in: Proceedings of the ACM Web Conference 2022, 2022, pp. 3217–3228.

[27] T. Gangopadhyay, S. Y. Tan, Z. Jiang, R. Meng, S. Sarkar, Spatiotemporal attention for multivariate time series prediction and interpretation, in: ICASSP 2021-2021 IEEE international conference on acoustics, speech and signal processing (ICASSP), IEEE, 2021, pp. 3560–3564.

[28] Z. Chang, X. Zhang, S. Wang, S. Ma, W. Gao, Stam: A spatiotemporal attention based memory for video prediction, IEEE Transactions on Multimedia 25 (2022) 2354–2367.

[29] T. Akiba, S. Sano, T. Yanase, T. Ohta, M. Koyama, Optuna: A next-generation hyperparameter optimization framework, in: Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, 2019, pp. 2623–2631.

[30] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, Q. Hu, Eca-net: Efficient channel attention for deep convolutional neural networks, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020, pp. 11534–11542.

[31] D. Chansong, S. Supratid, Impacts of kernel size on different resized images in object recognition based on convolutional neural network, in: 2021 9th international electrical engineering congress (iEECON), IEEE, 2021, pp. 448–451.

[32] M.-I. Georgescu, R. T. Ionescu, A.-I. Miron, O. Savencu, N.-C. Ristea, N. Verga, F. S. Khan, Multimodal multi-head convolutional attention with various kernel sizes for medical image super-resolution, in: Proceedings of the IEEE/CVF winter conference on applications of computer vision, 2023, pp. 2195–2205.

[33] M. Tan, Q. V. Le, Mixconv: Mixed depthwise convolutional kernels, arXiv preprint arXiv:1907.09595 (2019).

[34] T. Liu, M. Chen, M. Zhou, S. S. Du, E. Zhou, T. Zhao, Towards understanding the importance of shortcut connections in residual networks, Advances in neural information processing systems 32 (2019).

[35] M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, C. Pal, The importance of skip connections in biomedical image segmentation, in: International workshop on deep learning in medical image analysis, international workshop on large-scale annotation of biomedical data and expert label synthesis, Springer, 2016, pp. 179–187.

[36] C. Wei, S. Kakade, T. Ma, The implicit and explicit regularization effects of dropout, in: International conference on machine learning, PMLR, 2020, pp. 10181–10192.

[37] S. Woo, J. Park, J.-Y. Lee, I. S. Kweon, Cbam: Convolutional block attention module, in: Proceedings of the European conference on computer vision (ECCV), 2018, pp. 3–19.

[38] D. Lasseux, F. J. Valdés-Parada, B. D. Wood, Recent developments in upscaling and characterization of flow and transport in porous media (2021).

[39] A. Szymkiewicz, A. Szymkiewicz, Upscaling from darcy scale to field scale, Modelling Water Flow in Unsaturated Porous Media: Accounting for Nonlinear Permeability and Material Heterogeneity (2013) 139–175.

[40] X. Zhang, F. Ma, S. Yin, C. D. Wallace, M. R. Soltanian, Z. Dai, R. W. Ritzi, Z. Ma, C. Zhan, X. Lü, Application of upscaling methods for fluid flow and mass transport in multi-scale heterogeneous media: A critical review, Applied energy 303 (2021) 117603.

[41] P. Horgue, R. Guibert, H. Gross, P. Creux, G. Debenest, Efficiency of a two-step upscaling method for permeability evaluation at darcy and pore scales, Computational Geosciences 19 (2015) 1159–1169.