Riverbed Estimation Using Locally-Structured Unitary Network

Seiyu HITOMI*, Hiroyasu YASUDA*, Kiyoshi HAYASAKA*and Shogo MURAMATSU*
* Niigata University, Niigata, Japan
f24c083a@mail.cc.niigata-u.ac.jp, hiro@gs.niigata-u.ac.jp,
hayasaka@hep.sc.niigata-u.ac.jp, shogo@eng.niigata-u.ac.jp

***Abstract*—This paper presents a novel method for riverbed state estimation by integrating a Locally Structured Unitary Network (LSUN) with Dynamic Mode Decomposition (DMD). Prior work by Kobayashi *et al.* employed Multiresolution Convolutional Sparse Coded DMD (MR-CSC-DMD) to model the time evolution of river dynamics. However, MR-CSC-DMD suffers from high memory usage and computational complexity due to the use of redundant filter banks required for sparse approximation, which entails mapping to even higher-dimensional spaces. In contrast, the proposed approach leverages LSUN to directly capture local coordinate systems of the data manifold structural sparsity constraints. To validate its effectiveness, this study conducted experiments using time-series data obtained from a physical river model. The experimental results demonstrate that the proposed method achieves superior performance to MR-CSC-DMD.**

I. Introduction

Recent advancements in sensing, networking, and computational technologies have significantly contributed to the analysis of complex dynamic phenomena. In particular, data-driven approaches for modeling dynamical systems from observational data are being actively developed in fields such as engineering and physics [1], [2]. This study focuses on Extended Dynamic Mode Decomposition (EDMD) [3], [4], a technique that enhances Dynamic Mode Decomposition (DMD) [5] by introducing a dictionary-based formulation, thereby enabling the linear approximation of nonlinear dynamical systems. Unlike traditional process-driven methods, EDMD does not require first-principle models, making it suitable for capturing features directly from empirical data. This flexibility allows EDMD to be applied to complex, nonlinear phenomena for which analytical models are difficult to obtain. In this study, we apply EDMD to the estimation of riverbed states.

The growing frequency and intensity of rainfall due to climate change have increased the risk of water-related disasters. Severe weather events resulted in river flooding, causing substantial damage to infrastructure and loss of human life [6], [7]. Accordingly, there is an urgent need for effective countermeasures to mitigate the impact of such disasters. The primary causes of river flooding include levee overtopping due to rising water levels and breaches caused by channel deformation. Accurate prediction and control of channel morphology are therefore essential for flood prevention. However, despite longstanding research efforts, mathematical models capable of providing reliable early warnings and control remain underdeveloped. Kaneko *et al.* [8] introduced a data-driven approach based on Convolutional Sparse Coded Dynamic Mode Decomposition (CSC-DMD) for analyzing riverbed dynamics. Subsequently, building on this work, Kobayashi *et al.* [9] developed Multiresolution Convolutional Sparse Coded DMD (MR-CSC-DMD), which uses Non-Separable Oversampled Lapped Transform (NSOLT) [10] to extract local structures from riverbed data via convolutional sparse coding. Despite of its effectiveness, MR-CSC-DMD relies on redundant dictionaries to capture local coordinate systems, resulting in increased dimensionality, memory usage, and computational burden—particularly problematic for high-dimensional natural data such as riverbed profiles.

To address these challenges, we propose a riverbed state estimation framework based on the Locally Structured Unitary Network (LSUN) [11]. LSUN provides a shift-variant linear representation capable of directly sampling the local coordinate systems of the data manifold as a bundle of the tangent spaces, thereby offering a promising solution for dynamics modeling.

In the study by Yasas *et al.*[11], LSUN was utilized as a preprocessing step for DMD in dynamics modeling, enabling dimensionality reduction, computational cost savings, and the extraction of essential features from data by assisting Proper Orthogonal Decomposition (POD). However, the extension to incorporate multi-resolution representation and the application to real-world data remained as open challenges. In this paper, we propose applying LSUN with multi-resolution representation to dynamics modeling for riverbed state estimation. The contributions of this study are summarized as follows:

- Introduction of multi-resolution representation into LSUN,
- Dynamics modeling using LSUN with delayed components incorporated into the state vector, and
- Implementation of riverbed state estimation based on LSUN-driven dynamics modeling.

We validate the proposed method through estimation experiments using riverbed time-series data obtained from a physical river model and measured by Stream Tomography (ST) [12]. A portion of the dataset is used to train the riverbed state predictor, and the remainder is used for testing. The experimental results demonstrate that the proposed method achieves superior estimation performance to MR-CSC-DMD.

979-8-3315-7206-8/25/$31.00

II. Overview of Data-Driven Dynamics Analysis

This section reviews MR-CSC-DMD as a previous work for data-driven analysis methods of riverbed time evolution.

A. CSC-DMD

CSC-DMD consists of two main components: feature extraction based on convolutional sparse representation, and time-evolution modeling based on EDMD[8].

Given a time-series dataset $\{\mathbf{x}_k\}_{k=0}^{N-1}$ in the state space $\mathcal{M} \subset \mathbb{R}^m$, CSC-DMD nonlinearly maps the data to a corresponding feature sequence $\{\mathbf{y}_k\}_{k=0}^{N-1}$ in a feature space $\tilde{\mathcal{F}} \subseteq \mathbb{R}^L$ via sparse approximation. For each state $\mathbf{x}_k$, the corresponding sparse coefficient $\mathbf{y}_k$ is obtained by solving the following optimization problem:

$$\{\hat{\mathbf{D}}, \{\hat{\mathbf{y}}_k\}\} = \arg\min_{\mathbf{D}, \{\mathbf{y}_k\}} J(\mathbf{D}, \{\mathbf{y}_k\} | \{\mathbf{x}_k\}), \tag{1}$$

where $\mathbf{D}$ denotes a synthesis dictionary and

$$J(\mathbf{D}, \{\mathbf{y}_k\} | \{\mathbf{x}_k\}) := \sum_{k=0}^{N-1} \left(\frac{1}{2} \|\mathbf{x}_k - \mathbf{D}\mathbf{y}_k\|_2^2 + \lambda \rho(\mathbf{y}_k) \right). \tag{2}$$

$J(\cdot|\cdot)$ evaluates the fidelity of the dictionary $\mathbf{D}$ and the sparse coefficient sequence $\{\mathbf{y}_k\}$ to the input data $\{\mathbf{x}_k\}$. This is a joint optimization problem with respect to both the dictionary and the coefficients, and is typically solved by alternately updating the sparse codes and dictionary in an iterative fashion.

In CSC-DMD, the synthesis dictionary $\mathbf{D}$ is modeled as a parameterized convolutional multidimensional filter bank $\mathbf{D}_{\boldsymbol{\Theta}}$. The dictionary optimization is formulated as:

$$\hat{\mathbf{D}} = \mathbf{D}_{\hat{\boldsymbol{\Theta}}}, \quad \hat{\boldsymbol{\Theta}} = \arg\min_{\boldsymbol{\Theta}} \frac{1}{2} \sum_{k=0}^{N-1} \|\mathbf{x}_k - \mathbf{D}_{\boldsymbol{\Theta}} \hat{\mathbf{y}}_k\|_2^2. \tag{3}$$

Using the learned convolutional dictionary $\hat{\mathbf{D}}$, the nonlinear mapping $\boldsymbol{\Psi} : \mathcal{M} \to \tilde{\mathcal{F}}$ from state space to feature space is established.

We assume that in the feature space $\tilde{\mathcal{F}}$, the nonlinear evolution $F(\cdot)$ can be approximated by a linear operator $\mathbf{K} \in \mathbb{R}^{L \times L}$. Accordingly, the temporal evolution of the feature vector $\mathbf{y}_k$ is expressed as:

$$\mathbf{y}_k \approx \mathbf{K}\Psi(\mathbf{x}_{k-1}) = \mathbf{K}\mathbf{y}_{k-1}. \tag{4}$$

Using the EDMD framework, the matrix $\mathbf{K}$ can be estimated in a data-driven manner. The procedure for estimating $\mathbf{K}$ is as follows:

Given the feature vectors $\mathbf{y}_k$, we define the data matrix $\mathbf{Y}_{k:\ell}$ for $k \leq \ell \leq N-1$ as:

$$\mathbf{Y}_{k:\ell} := \begin{bmatrix} \mathbf{y}_k & \mathbf{y}_{k+1} & \cdots & \mathbf{y}_\ell \end{bmatrix} \in \mathbb{R}^{L \times (\ell - k + 1)}. \tag{5}$$

1) Perform singular value decomposition (SVD) on $\mathbf{Y}_{0:N-2}$ to obtain $\mathbf{U}_r \in \mathbb{C}^{L \times r}$, $\boldsymbol{\Sigma}_r \in \mathbb{C}^{r \times r}$, and $\mathbf{V}_r \in \mathbb{C}^{(N-1) \times r}$, where $r \leq \operatorname{rank}(\mathbf{Y}_{0:N-2})$.

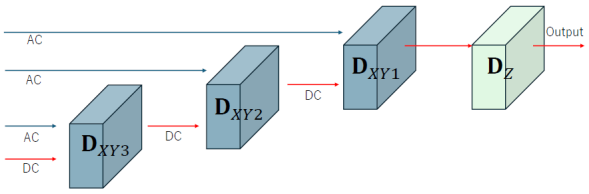


Fig. 1: Schematic diagram of multiresolution representation, illustrating an example of a 3-level tree structure. DC and AC denote the direct current and and alternating components, respectively.

2) Project $\mathbf{K}$ onto $\mathbf{U}_r$ to obtain:

$$\check{\mathbf{K}} = \mathbf{U}_r^{\mathsf{H}} \mathbf{Y}_{2:N-1} \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1}, \tag{6}$$

where H denotes the Hermitian transpose.

3) Compute the eigendecomposition $\check{\mathbf{K}}\mathbf{W} = \mathbf{W}\boldsymbol{\Lambda}$, where $\mathbf{W} \in \mathbb{C}^{r \times r}$ and $\boldsymbol{\Lambda} \in \mathbb{C}^{r \times r}$ is a diagonal matrix containing the eigenvalues of $\check{\mathbf{K}}$.

4) Compute the dynamic modes of $\mathbf{K}$ as:

$$\boldsymbol{\Phi} = \mathbf{Y}_{1:N-1} \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1} \mathbf{W}. \tag{7}$$

5) The operator $\mathbf{K}$ is then approximated as:

$$\mathbf{K} \approx \boldsymbol{\Phi}\boldsymbol{\Lambda}\boldsymbol{\Phi}^{\dagger}, \tag{8}$$

where $\dagger$ denotes the Moore–Penrose pseudoinverse. Based on this formulation, the time evolution of the feature vector $\mathbf{y}_k \in \tilde{\mathcal{F}}$ is approximated by

$$\mathbf{y}_k \approx \boldsymbol{\Phi}\boldsymbol{\Lambda}^k \mathbf{b}_0, \quad \mathbf{b}_0 = \boldsymbol{\Phi}^{\dagger} \mathbf{y}_0. \tag{9}$$

Using the feature vector $\mathbf{y}_k \in \mathbb{R}^L$, the state vector $\mathbf{x}_k \in \mathbb{R}^m$ is reconstructed as:

$$\mathbf{x}_k \approx \mathbf{V}\boldsymbol{\Lambda}^k \mathbf{b}_0, \tag{10}$$

where $\mathbf{V} = \hat{\mathbf{D}}\boldsymbol{\Phi}$. The continuous-time formulation is given by

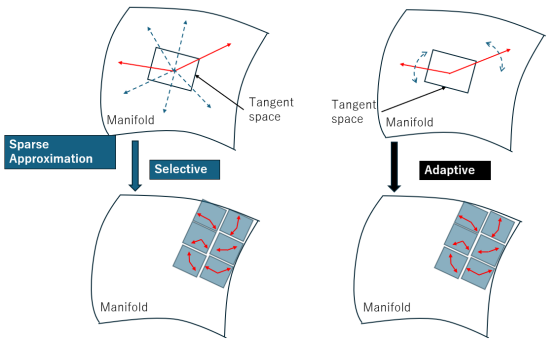
$$\mathbf{x}(t) \approx \mathbf{V} e^{\boldsymbol{\Omega} t} \mathbf{b}_0, \quad \boldsymbol{\Omega} = \frac{\ln \boldsymbol{\Lambda}}{\Delta t}. \tag{11}$$

In this way, CSC-DMD combines sparse feature representation based on a convolutional dictionary and time evolution modeling via EDMD to effectively capture the dynamics of nonlinear systems.

B. MR-CSC-DMD

MR-CSC-DMD, an extension of CSC-DMD that incorporates multiresolution representation in the spatial domain[9]. MR-CSC-DMD utilizes a $(2+1)$-dimensional synthesis dictionary that separates spatial and temporal (stack) directions. This synthesis dictionary is constructed as a tensor product of a two-dimensional non-separable dictionary $\mathbf{D}_{\mathrm{XY}}$ and a one-dimensional non-overlapping dictionary $\mathbf{D}_{\mathrm{Z}}$, denoted as $\mathbf{D}_{\mathrm{XY}} \otimes \mathbf{D}_{\mathrm{Z}}$.

Despite of the performance improvement , MR-CSC-DMD faces challenges in high-dimensional settings. The redundant



(a) (b)

Fig. 2: Illustration of two methods for capturing tangent spaces. Each arrow indicates a coordinate axis corresponding to a filter. Solid lines represent used coordinate axes, while dotted lines represent ignored coordinate axes. (a) and (b) show a method using sparse modeling and one using LSUN, respectively.

filter banks increase computational and memory demands due to the necessity of higher-dimensional operations.

III. Riverbed State Estimation Using LSUN

In this section, we propose to introduce LSUN for riverbed state estimation. LSUN extends the concept of block principal component analysis (PCA) to overlapping local structures, enabling data-driven learning of tangent spaces while preserving shift-variance in local coordinates [11][13]. LSUN employs an adaptive mechanism instead of convolutional processing with a selection mechanism such as sparse approximation. Fig. 2 compares the two different approaches for capturing tangent spaces.

Yasas *et al.* proposed a dynamics modeling approach that combines LSUN as a preprocessing step for DMD, using it to support POD. This approach allows dimensionality reduction, computational cost reduction, and extraction of essential features from data. LSUN still has potential for better performance by introducing multiresolution representations.

In the proposed method, we adopt a state vector formulation that incorporates multi-resolution representation and delay embedding technique, both of which were shown to be advantageous for riverbed state estimation in the CSC-DMD work by Kobayashi *et al.*[9][14].

A. Training of LSUN

We employ LSUN with multiresolution representation as a two-dimensional non-separable dictionary $\hat{\mathbf{D}}_{\mathrm{XY}}$ with a tree-structure of hierarchy level ℓ. The first goal is to learn a unitary synthesis dictionary $\hat{\mathbf{D}}_{\mathrm{XY}}$ that concentrates the input energy $\|\mathbf{x}\|_2^2$ of the input data $\mathbf{x} \in \mathbb{R}^m$ in a lower-dimensional subspace. We structurally force LSUN to have the No-DC leakage property. The extracted DC components are further analyzed by another LSUN in next level.

The LSUN design problem is formulated as follows:

$$\{\hat{\boldsymbol{\Theta}}_\ell\}_\ell = \underset{\{\boldsymbol{\Theta}_\ell\}_\ell}{\operatorname{argmax}} \sum_k \left[\sum_{\ell=1}^{\tau-1} \left\| \boldsymbol{\Gamma}_{\mathrm{AC}}^{(\ell)} \mathbf{D}_{\mathrm{XY}\boldsymbol{\Theta}_\ell}^{(\ell)\top} \mathbf{v}_k^{(\ell)} \right\|_2^2 + \left\| \boldsymbol{\Gamma}^{(\tau)} \mathbf{D}_{\mathrm{XY}\boldsymbol{\Theta}_\tau}^{(\tau)\top} \mathbf{v}_k^{(\tau)} \right\|_2^2 \right], \tag{12}$$

where $\boldsymbol{\Gamma}^{(\ell)} \in \mathbb{R}^{B^{(\ell)}M \times B^{(\ell)}M}$ is the channel extraction matrix with elements set to 1 for BK entries and 0 for the remaining $B(M-K)$ entries. Let $M = M_{\mathrm{X}} \times M_{\mathrm{Y}}$, and the number of blocks at tree level ℓ is $B^{(\ell)}$. $\mathbf{D}_{\mathrm{XY}\boldsymbol{\Theta}}^{(\ell)} \in \mathbb{R}^{B^{(\ell)}M \times B^{(\ell)}M}$ denotes the LSUN obtained through training in the ℓ-level tree. Note that LSUN structurally satisfies $\mathbf{D}_{\mathrm{XY}\boldsymbol{\Theta}}^{(\ell)} \mathbf{D}_{\mathrm{XY}\boldsymbol{\Theta}}^{(\ell)\top} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix. For each sample number k, the input to the first layer is denoted as $\mathbf{v}_k^{(1)} := \mathbf{x}_k$, and the input to each subsequent layer is recursively defined by propagating the DC components extracted from the previous layer:

$$\mathbf{v}_k^{(\ell+1)} := \boldsymbol{\Gamma}_{\mathrm{DC}}^{(\ell)} \mathbf{D}_{\mathrm{XY}\boldsymbol{\Theta}}^{(\ell)\top} \mathbf{v}_k^{(\ell)}. \tag{13}$$

$\boldsymbol{\Gamma}^{(\ell)}$ is split into two matrices: $\boldsymbol{\Gamma}_{\mathrm{DC}}^{(\ell)} \in \{0,1\}^{B^{(\ell)} \times B^{(\tau)}}$, which extracts one DC channel per block, and $\boldsymbol{\Gamma}_{\mathrm{AC}}^{(\ell)} \in \{0,1\}^{B^{(\ell)}(M-1) \times B^{(\ell)}(M-1)}$, which extracts the remaining $(K-1)$ channels per block, representing the residual (AC) components. These matrices satisfy the following direct sum relation:

$$\boldsymbol{\Gamma}^{(\ell)} = \boldsymbol{\Gamma}_{\mathrm{DC}}^{(\ell)} \oplus \boldsymbol{\Gamma}_{\mathrm{AC}}^{(\ell)}. \tag{14}$$

Unlike the dictionary design of the conventional MR-CSC-DMD, the proposed approach does not require any redundant dictionary. Instead, it directly captures local coordinate systems, thereby addressing challenges related to memory consumption and computational cost associated with mapping to higher-dimensional representations.

B. LSUN+POD+DMD

We propose a novel approach for riverbed estimation using LSUN, POD and DMD. To estimate the riverbed state $\mathbf{x}_{\mathrm{b}_k}$, the state vector $\mathbf{x}_k$ is defined as

$$\mathbf{x}_k := \begin{pmatrix} \mathbf{x}_{\mathrm{b}_k} \\ \mathbf{x}_{\mathrm{b}_{k-1}} \end{pmatrix} \in \mathbb{R}^m, \tag{15}$$

where $\mathbf{x}_{\mathrm{b}_k} \in \mathbb{R}^{m_{\mathrm{b}}}$ is the riverbed state at time $t = k\Delta t$, a delay embedding is incorporated into the state vector $\mathbf{x}_{\mathrm{b}_k}$, leading to $m = 2m_{\mathrm{b}}$. Using the time-series data $\{\mathbf{x}_{\mathrm{b}_k}\}$, we train $\hat{\mathbf{D}}_{\mathrm{XY}} := \mathbf{D}_{\mathrm{XY}\hat{\boldsymbol{\Theta}}} \mathbf{S}_{B,K}^\top$. Here, $\mathbf{S}_{B,K} \in \{0,1\}^{BK \times BM}$ is a subsampling matrix, where the relation $\boldsymbol{\Gamma}_{B,K} = \mathbf{S}_{B,K}^\top \mathbf{S}_{B,K}$ holds. Using the learned $\hat{\mathbf{D}}_{\mathrm{XY}}$ and time-series data $\{\mathbf{x}_k\}_{k=0}^{N-1}$, the LSUN coefficient data matrix is defined as

$$\mathbf{Y}_0 := \hat{\mathbf{D}}_{\mathrm{XY}}^\top \begin{bmatrix} \mathbf{x}_0 & \mathbf{x}_1 & \cdots & \mathbf{x}_{N-2} \end{bmatrix}. \tag{16}$$

Then the leading r POD modes $\mathbf{U}_r$ is obtained by SVD for LSUN coefficient data matrix $\mathbf{Y}_0$. $\mathbf{U}_r$ takes the place of $\mathbf{D}_{\mathrm{Z}}^\top$

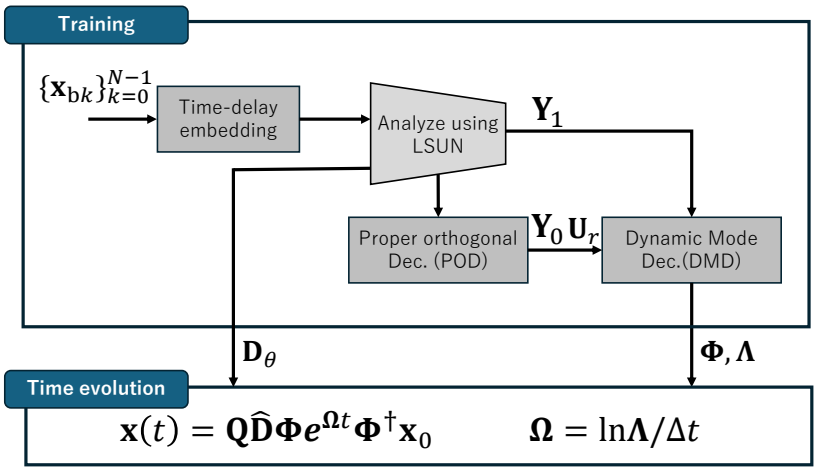


Fig. 3: Overview of the proposed method. $\{\mathbf{x}_{\mathrm{b}k}\}$ is the riverbed time-series data. $\hat{\mathbf{D}}_{\mathrm{XY}}$ is the unitary synthesis dictionary. $\mathbf{U}_r$ denotes the principal POD modes.$\mathbf{Y}_0$ and $\mathbf{Y}_1$ are LSUN coefficient data matrix.$\boldsymbol{\Phi}$ is a dynamic mode. $\boldsymbol{\Omega}$ and $\boldsymbol{\Lambda}$ are eigenvalues. $\mathbf{Q}$ is defined as $\mathbf{Q} := [\mathbf{I} \quad \mathbf{O}]$.

in the MR-CSC-DMD framework. The matrix $\hat{\mathbf{K}}$ is obtained by solving the optimization problem

$$\hat{\mathbf{K}} = \arg\min_{\mathbf{K}} \left\| \mathbf{Y}_1 - \mathbf{U}_r^{\top}\mathbf{K}\mathbf{U}_r\mathbf{Y}_0 \right\|_F^2, \tag{17}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Also,

$$\mathbf{Y}_1 := \hat{\mathbf{D}}_{\mathrm{XY}}^{\top}\left[\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_{N-1}\right]. \tag{18}$$

Then, by applying the EDMD framework, the dynamic modes $\boldsymbol{\Phi}$ and eigenvalues $\boldsymbol{\Lambda}$ of $\hat{\mathbf{K}}$ are obtained. The time evolution is expressed as

$$\mathbf{x}(t) = \mathbf{Q}\mathbf{V}e^{\boldsymbol{\Omega}t}\mathbf{b}_0, \tag{19}$$

where $\mathbf{V} = \hat{\mathbf{D}}_{\mathrm{XY}}\boldsymbol{\Phi}$ and $\mathbf{b}_0 = \boldsymbol{\Phi}^{\dagger}\mathbf{x}_0$ and $\mathbf{Q}$ is the matrix that extracts $\mathbf{x}_{\mathrm{b}_k}$ from the vector $\mathbf{x}_k$, and is defined as $\mathbf{Q} := [\mathbf{I} \quad \mathbf{O}]$, where $\mathbf{I}$ and $\mathbf{O}$ are the identity and zero matrices, respectively.

An overview of the proposed method for deriving the time evolution equation is shown in Fig. 3.

IV. Performance Evaluation

To verify the effectiveness of the proposed method, we evaluate its performance through experiments on riverbed state estimation. In this study, we utilize an experimental river setup and ST developed by Moteki *et al.* [12], to simultaneously measure the shapes of the water surface and the riverbed. ST is a mechanical device that scans the reflective positions of flowing water and moving sediment using a sheet laser and two cameras facing each other with the laser sheet between them. The experimental river model has a length of 12 meters, a width of 0.45 meters, and a channel slope of gradient 1/200.

The data were sampled every 10 minutes from 100 to 440 minutes after the start of inflow. The data from 100 to 250 minutes were used for training data $\{\mathbf{x}_k\}_{k=0}^{N-1}$, and the data from 260 to 440 minutes were used for testing. The effectiveness of the proposed method is evaluated using experimental data with LSUN[1] configuration parameters; block size [M_{X}, M_{Y}] = [2, 4], the tree level $\tau = 4$, the number of coefficients per block $K^{(\ell)}$ was set to $[K^{(1)}, K^{(2)}, K^{(3)}, K^{(4)}] = [1, 1, 1, 4]$ corresponding to each tree level ℓ, which match to the number of non-zero components per block in the sparse coefficients for MR-CSC-DMD. The number of overlapping blocks $[J_{\mathrm{X}}, J_{\mathrm{Y}}]$ is varied for $J_{\mathrm{X}} = J_{\mathrm{Y}} \in \{1, 3, 5\}$ and the number of POD modes r is varied from 8 to 14.

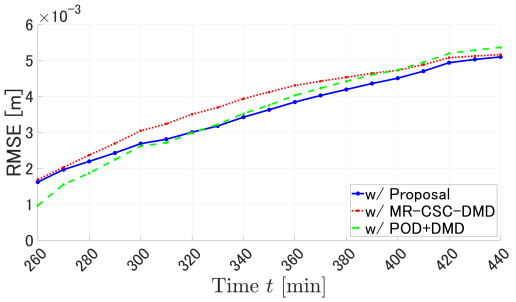


Fig. 4: Comparison between the conventional method and the proposed method for riverbed estimation.

We evaluate the performance of the proposed method using the root mean square error (RMSE) between estimated and observed data.

A. Estimation results

We compare the estimation performance of using our proposed method against the previous MR-CSC-DMD and

[1]GitHub TanSacNet Project https://github.com/msiplab/TanSacNet

TABLE I: Comparison of state vector and processing time in the experiment. $\mathbf{y}_k = \hat{\mathbf{D}}_{\mathrm{XY}}^{\top}\mathbf{x}_k$ is the size of the feature vector.

	Proposal	POD+DMD	MR-CSC-DMD
$\mathbf{y}_k$	48 × 1	24576 × 1	126384 × 1
Processing time [sec]	0.25	0.10	0.53

TABLE II: Riverbed estimation results of proposal for varying the numbers of POD modes r and overlapping blocks $[J_{\mathrm{X}}, J_{\mathrm{Y}}]$. RMSE represents the average over the entire estimation period. The block size and the tree level are fixed to $[M_{\mathrm{X}}, M_{\mathrm{Y}}] = [2, 4]$ and $\tau = 4$, respectively.

r	$[J_{\mathrm{X}}, J_{\mathrm{Y}}]$	RMSE [m]
14 (full modes)	[1,1]	4.3×10^{-3}
14 (full modes)	[3,3]	3.9×10^{-3}
14 (full modes)	[5,5]	4.0×10^{-3}
7	[3,3]	3.6×10^{-3}
9	[3,3]	$\mathbf{3.6 \times 10^{-3}}$
11	[3,3]	3.9×10^{-3}

POD+DMD method. The former method shows the best performance by varying the number of overlapping blocks and the regularization parameter, using the configuration in reference [9] as a baseline. The latter method performs dimensionality reduction in the temporal direction by projecting the data onto the POD basis without applying LSUN. Table I shows the size comparison of the state vectors $\mathbf{y}_k = \hat{\mathbf{D}}_{\mathrm{XY}}^{\top}\mathbf{x}_k$ and the average processing time per frame required to estimate the riverbed state through the derived time evolution equation. The computational specifications and simulation tools used to measure the processing time were Intel Core i5-14400F CPU, 32GB RAM, and MATLAB R2022b. Fig. 4 shows the RMSE results. The experimental results demonstrate that the proposed method enables dimension reduction of the dynamics while achieving more accurate riverbed state estimation than the others. The processing time comparison shows that the proposed method achieves improved performance and faster processing speed compared to MR-CSC-DMD, although it is slower than POD+DMD. Note that the MR-CSC-DMD framework adopts a fixed Discrete Cosine Transform (DCT) in $\mathbf{D}_{\mathrm{Z}}$, as in reference [9]. It can be replaced with POD, but this does not effect the dimensionality reduction of features. Regarding the computational cost of POD, the proposed method significantly reduces memory usage and computation power because the analysis is applied to the low-dimensional coefficient matrix $\mathbf{Y}_0$.

Table II summarizes the results obtained by varying the number of overlapping blocks in LSUN and the number of POD modes used. The block size is fixed to $[M_{\mathrm{X}}, M_{\mathrm{Y}}] = [2, 4]$, and the tree level is set to $\tau = 4$. From the table results, we confirmed that the use of spatial overlapping in LSUN improves the performance, and that POD-based mode reduction also contributes to performance enhancement.

Table III summarizes the results obtained by varying the tree level τ and applying delay embedding. The results confirm that both multiresolution representation and delay embedding contribute to the performance improvement, and that the best performance is achieved when both are applied. Fig. 5 shows the estimation results when multi-resolution representation is incorporated into LSUN and applying delay embedding. By increasing the tree level τ, it becomes possible to capture large-scale changes in the river, confirming the effectiveness of the method for long-term estimation.

TABLE III: Riverbed estimation results of proposal for varying tree level τ and delay embedding. RMSE represents the average over the entire estimation period.

Tree level τ	Delay embedding	RMSE [m]
1	No	4.0×10^{-3}
1	Yes	3.7×10^{-3}
4	No	3.9×10^{-3}
4	Yes	$\mathbf{3.6 \times 10^{-3}}$

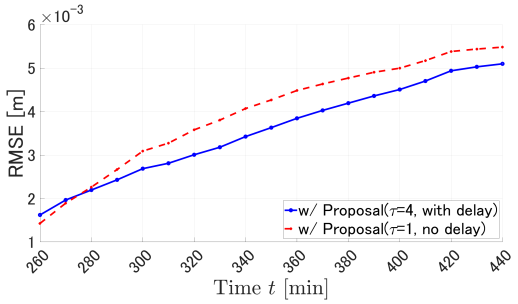


Fig. 5: Effect of multiresolution representation and delay embedding. Comparison of tree level $\tau = 1$ and $\tau = 4$.

V. Conclusion

In this paper, we proposed riverbed state estimation using LSUN, POD and DMD, and evaluated its performance through estimation experiments. The proposed method was assessed by comparing in terms of the RMSE and the size of the feature vector sequence analyzed by DMD, confirming its effectiveness for riverbed state estimation. We also evaluated the riverbed state estimation performance by incorporating the multiresolution representation into LSUN and applying delay embedding.

By comparing the proposal with MR-CSC-DMD and POD+DMD, we verified that LSUN allows dimensionality reduction while providing more accurate riverbed state estimations in long-term estimation. Additionally, further improvement is observed by the multiresolution representation with LSUN and delay embedding.

Future work includes investigating performance variations by changing the tree level τ and the number of LSUN channels K.

Acknowledgment

This research was supported by JSPS KAKENHI Grant Numbers JP22H00512, JP24H00365, and JP21H04596.

References

[1] M. Cenedese, J. Axås, B. Bäuerlein, K. Avila, and G. Haller, "Data-driven modeling and prediction of non-linearizable dynamics via spectral submanifolds," *Nature Communications*, vol. 13, p. 872, 2022. DOI: 10.1038/s41467-022-28518-y.

[2] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of Fluid Mechanics*, vol. 656, pp. 5–28, 2010. DOI: 10.1017/S0022112010001217.

[3] M. O. Williams, I. G. Kevrekidis, and C. W. Rowley, "A data-driven approximation of the koopman operator: Extending dynamic mode decomposition," *Journal of Nonlinear Science*, vol. 25, no. 6, pp. 1307–1346, 2015. DOI: 10.1007/s00332-015-9258-5.

[4] H. Terao, S. Shirasaka, and H. Suzuki, "Extended dynamic mode decomposition with dictionary learning using neural ordinary differential equations," *Nonlinear Theory and Its Applications, IEICE*, vol. 12, no. 4, pp. 626–638, 2021. DOI: 10.1587/nolta.12.626.

[5] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of lipschitz–hankel type involving products of bessel functions," *Philosophical Transactions of the Royal Society of London Series A*, vol. 247, pp. 529–551, 1955. DOI: 10.1098/rsta.1955.0005.

[6] N. W. Arnell and S. N. Gosling, "The impacts of climate change on river flood risk at the global scale," *Climatic Change*, vol. 134, no. 3, pp. 387–401, 2016. DOI: 10.1007/s10584-014-1084-5.

[7] H. Alifu, Y. Hirabayashi, Y. Imada, and H. Shiogama, "Enhancement of river flooding due to global warming," *Scientific Reports*, vol. 12, no. 1, p. 20 687, 2022. DOI: 10.1038/s41598-022-25182-6.

[8] Y. Kaneko, S. Muramatsu, H. Yasuda, *et al.*, "Convolutional-sparse-coded dynamic mode decomposition and its application to river state estimation," in *2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 1872–1876. DOI: 10.1109/ICASSP.2019.8683848.

[9] E. Kobayashi, H. Yasuda, K. Hayasaka, Y. Otake, S. Ono, and S. Muramatsu, "Multi-resolution convolutional dictionary learning for riverbed dynamics modeling," in *2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 3897–3901. DOI: 10.1109/ICASSP49357.2023.10096452.

[10] S. Muramatsu, K. Furuya, and N. Yuki, "Multidimensional nonseparable oversampled lapped transforms: Theory and design," *IEEE Transactions on Signal Processing*, vol. 65, no. 5, pp. 1251–1264, 2017. DOI: 10.1109/TSP.2016.2633240.

[11] Y. Godage, E. Kobayashi, and S. Muramatsu, "Locally-structured unitary network," *APSIPA Transactions on Signal and Information Processing*, vol. 13, no. 1, 2024. DOI: 10.1561/116.00000308.

[12] D. Moteki, T. Murai, T. Hoshino, H. Yasuda, S. Muramatsu, and K. Hayasaka, "Capture method for digital twin of formation processes of sand bars," *Physics of Fluids34(034117)*, 2022. DOI: 10.1063/5.0085574.

[13] A. Liu, Y. Zhang, E. Gehan, and R. Clarke, "Block principal component analysis with application to gene microarray data classification," *Statistics in Medicine*, vol. 21, no. 22, pp. 3465–3474, 2002. DOI: 10.1002/sim.1263.

[14] D. J. Griffiths, *Introduction to Electrodynamics*, 5th. Cambridge University Press, 2023, ISBN: 978-1-009-39775-9.