Physics-Inspired Decomposition for Weak Low-Rank Spatiotemporal Completion in Sparse Crowdsensing

Mijia Zhang[1], En Wang[1,*], Wenbin Liu[1,*], Junwei Zhao[2], Yang Xu[3], Bo Yang[1], Jie Wu[2]

[1]College of Computer Science and Technology, Jilin University
[2]China Telecom Cloud Computing Research Institute
[3]College of Computer Science and Electronic Engineering, Hunan University

Presenter: Mijia Zhang

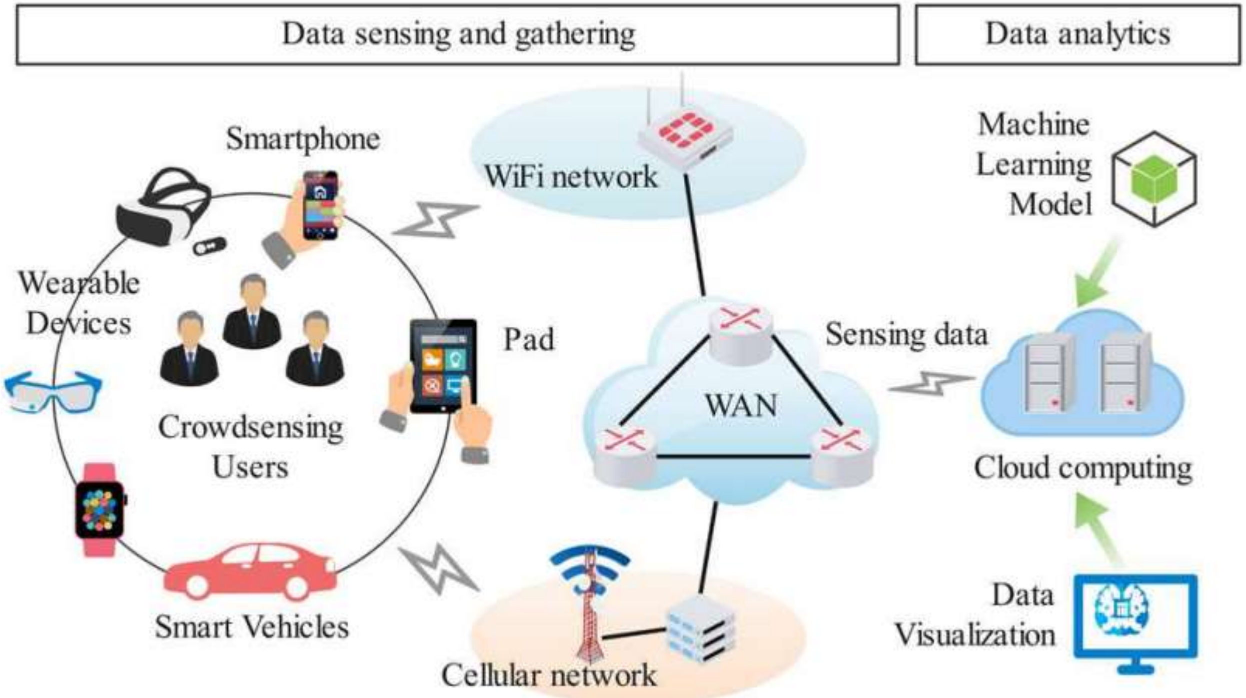
Outlines

1. Introduction and challenges
2. Contributions
3. Model
4. Experiments

1. **Introduction and challenges**
2. Contributions
3. Model
4. Experiments

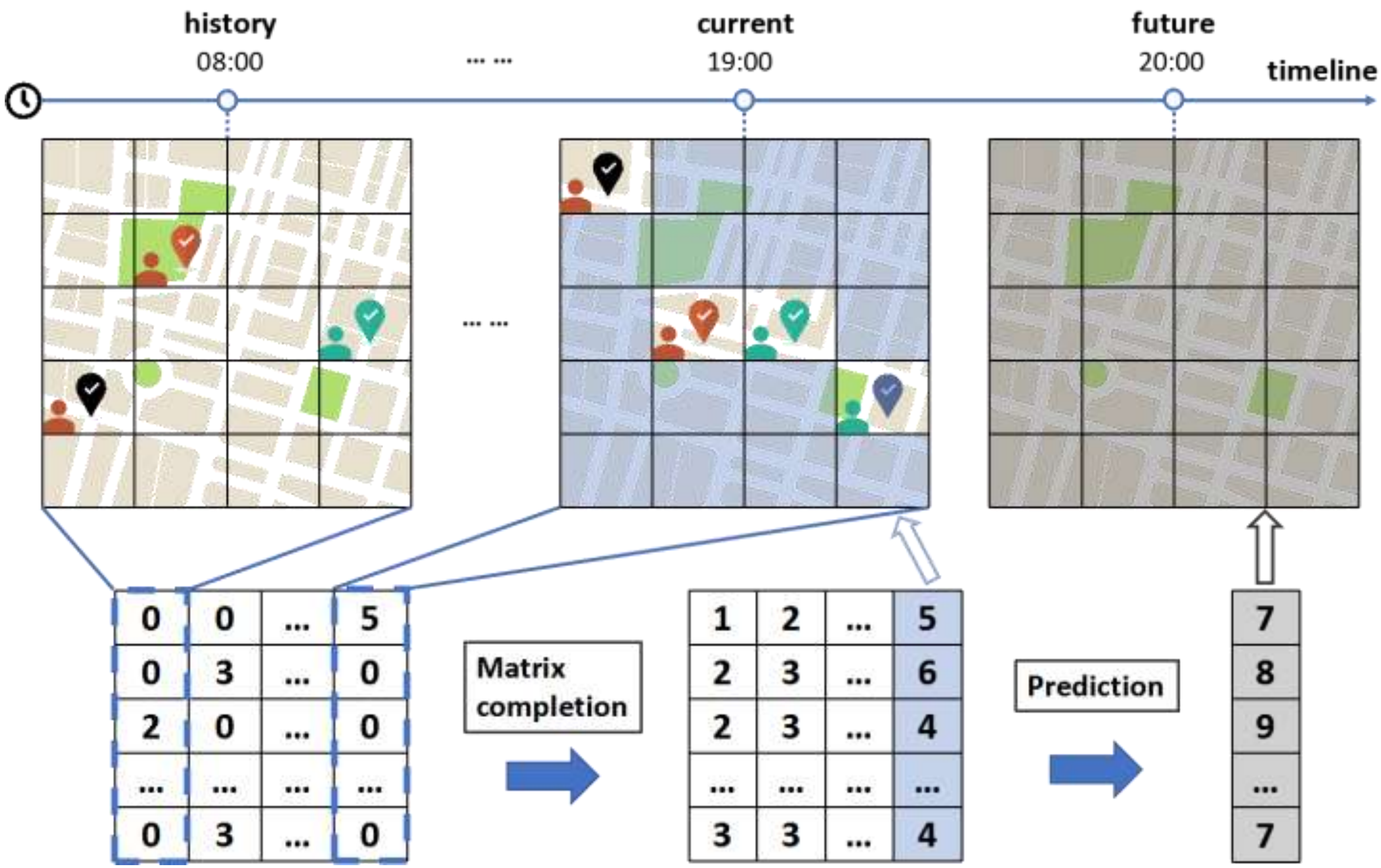
1. Introduction and challenges

❖ Mobile CrowdSensing (MCS)



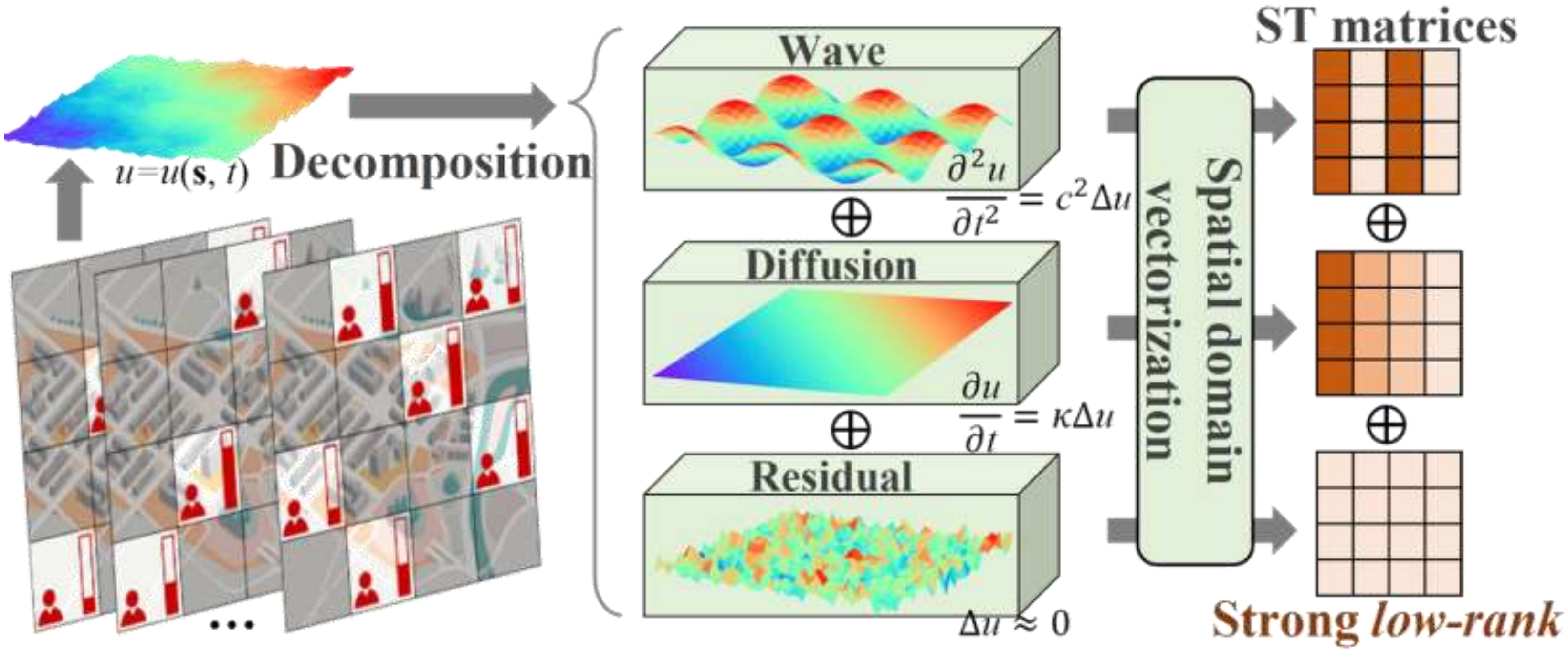
- Task publisher
- Server platform
- Task executor

- Sparse MCS


history
08:00
current
19:00
future
20:00
timeline
Matrix completion
Prediction
0 0 ... 5
0 3 ... 0
2 0 ... 0
0 3 ... 0
1 2 ... 5
2 3 ... 6
2 3 ... 4
3 3 ... 4
7
8
9
7

1. Introduction and challenges

Exemplary scenarios illustrating physics-inspired decomposition for interpretable spatiotemporal completion ideas and challenges in sparse crowdsensing.



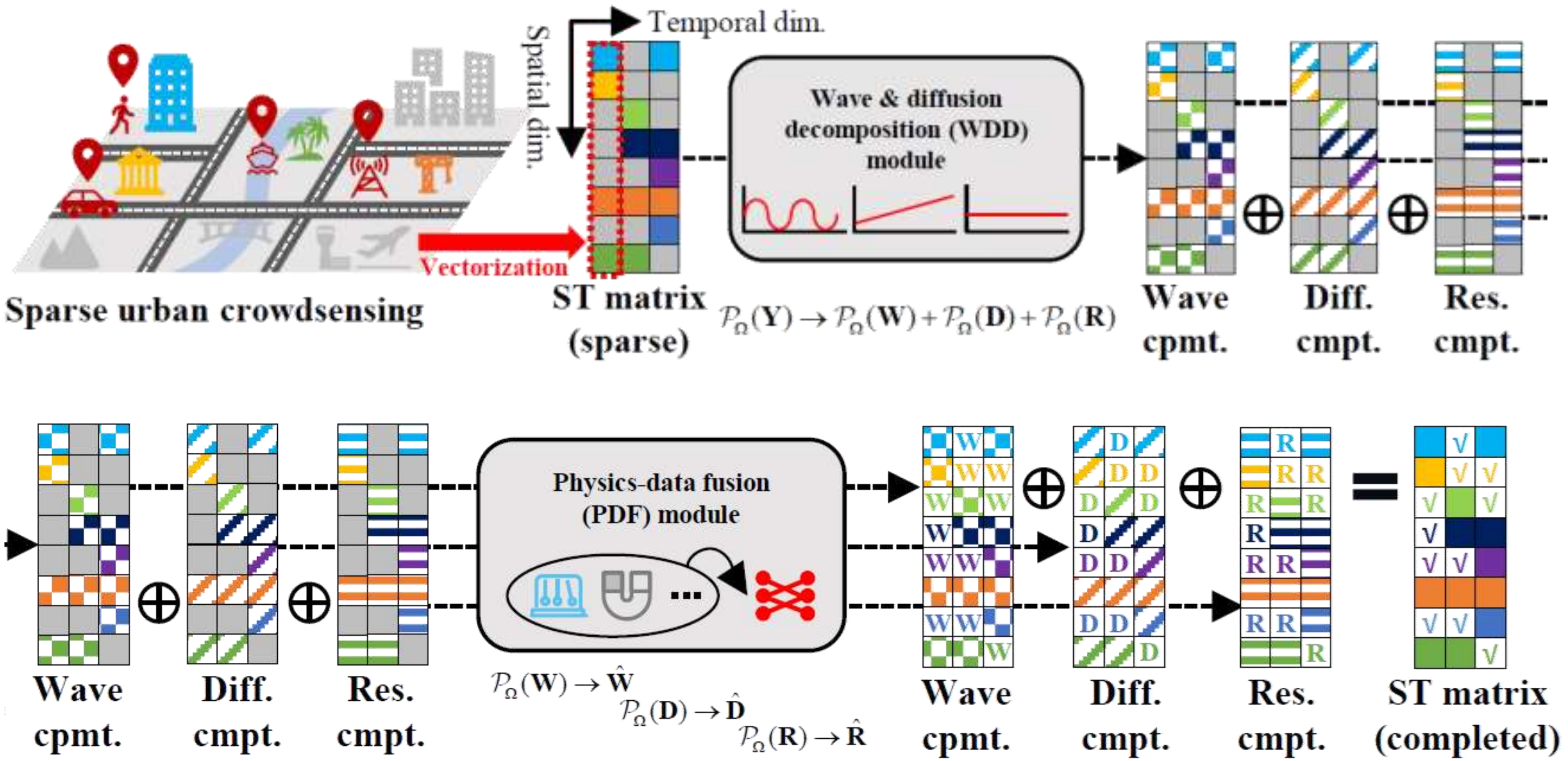
1. Introduction and challenges

Challenges

- the effective component decomposition method
- the subsequent processing after decomposition
- the spatiotemporal completion accuracy achieved by the physics-inspired method

1. Introduction and challenges

- The framework employs physics-inspired decomposition for interpretable spatiotemporal completion.



1. Introduction and challenges
2. **Contributions**
3. Model
4. Experiments

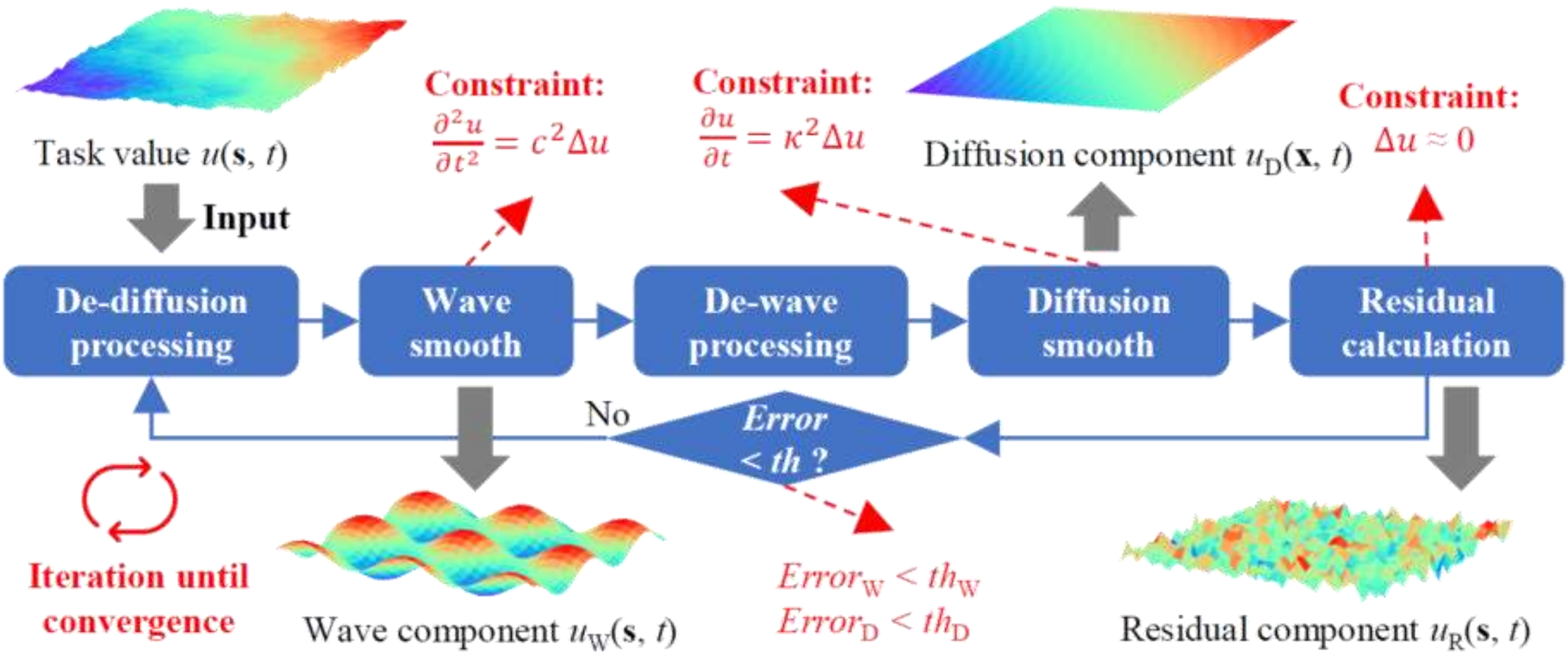
2. Contributions

- Our work has the following contributions:
  - We propose a novel spatiotemporal completion framework MCWDD, which uses PDEs to decompose data matrices into three physically interpretable components.
  - MCWDD employs analytical regularization to enforce strong low-rank properties in each component.
  - The combination of data and knowledge enabling robust completion for weak low-rank data.
  - Extensive experiments on synthetic and real-world data show MCWDD achieves higher completion accuracy.

1. Introduction and challenges
2. Contributions
3. **Model**
4. Experiments

3. Model

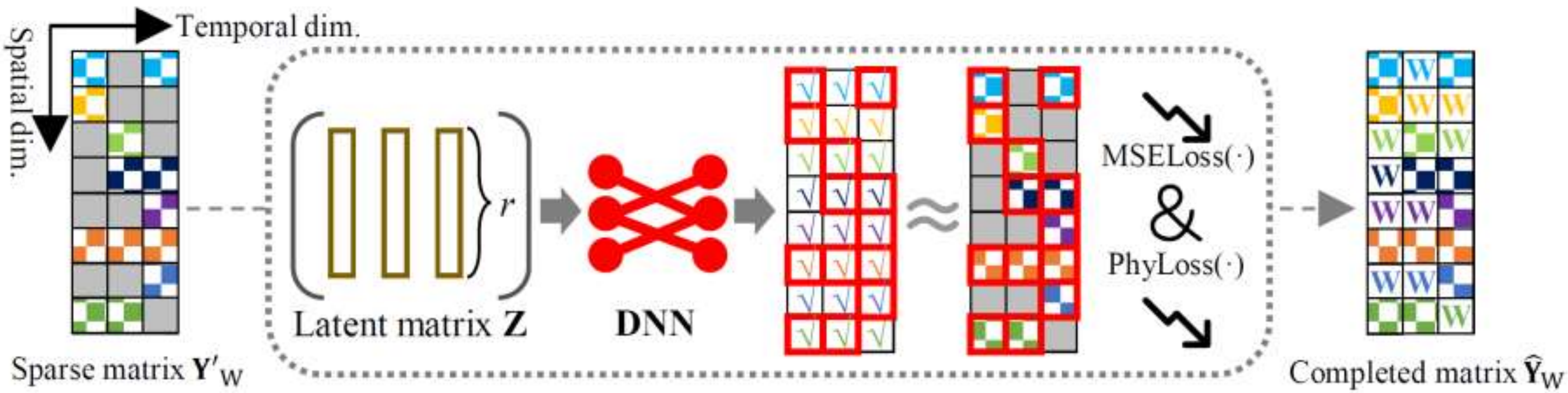
❖ Wave and Diffusion Decomposition Module



Loss Function: $\mathcal{L}_W(\mathbf{W}) = \frac{\partial^2 \mathbf{W}}{\partial t^2} - c^2 \Delta \mathbf{W}$ $\mathcal{L}_D(\mathbf{D}) = \frac{\partial \mathbf{D}}{\partial t} - \kappa \Delta \mathbf{D}$

3. Model

❖ Physics-Data Fusion Module



Deep Interaction Layers:

$$\mathbf{h}_j^{(0)} = \sigma_0(\mathbf{\Theta}^{(0)}\mathbf{q}_j), \qquad \mathbf{h}_j^{(1)} = \sigma_1(\mathbf{\Theta}^{(1)}\mathbf{h}_j^{(0)} + \mathbf{b}^{(1)}), \qquad \ldots \qquad \mathbf{h}_j^{(L)} = \sigma_L(\mathbf{\Theta}^{(L)}\mathbf{h}_j^{(L-1)} + \mathbf{b}^{(L)}), \qquad \hat{\mathbf{w}}_j = \sigma_{L+1}(\mathbf{\Theta}^{(L+1)}\mathbf{h}_j^{(L)} + \mathbf{b}^{(L+1)})$$

Loss Function :

$$\mathcal{L} = \underbrace{\sum_{j=1}^{N_t} ||\mathcal{P}_\Omega(\mathbf{w}_j - \hat{\mathbf{w}}_j)||^2}_{\text{MSE Loss}} + \underbrace{\lambda_1 \sum_{j=1}^{N_t} ||\mathbf{q}_j||^2 + \lambda_2 \sum_{l=0}^{L+1} ||\mathbf{\Theta}^{(l)}||^2 + \lambda_3 \sum_{l=1}^{L+1} ||\mathbf{b}^{(l)}||^2}_{\text{Regularization Penalty}} + \underbrace{\lambda_4 ||\mathcal{L}_W(\mathbf{W})||}_{\text{Physics Loss}},$$

1. Introduction and challenges
2. Contributions
3. Model
4. **Experiments**

4. Experiments

❖ Statistics of two evaluation datasets.

TABLE I: Statistics of the **U-Air** dataset.

Country/Region - City	Chinese mainland - Beijing	
# Location	36 (each with 1 km $\times$ 1 km)	
# Time step	264 (each with an hour)	
Data [Unit]	$PM_{2.5}$ [$\mu g/m^3$]	PM_{10} [$\mu g/m^3$]
Mean $\pm$ Std.	79.11 ± 81.21	63.12 ± 48.56

TABLE II: Statistics of the **Sensor-Scope** dataset.

Country/Region - City	Switzerland - Lausanne	
# Location	57 (each with 30 m $\times$ 50 m)	
# Time step	336 (each with half an hour)	
Data [Unit]	Temperature [$^\circ$C]	Humidity [%]
Mean $\pm$ Std.	6.04 ± 1.87	84.52 ± 6.32

❖ Baselines

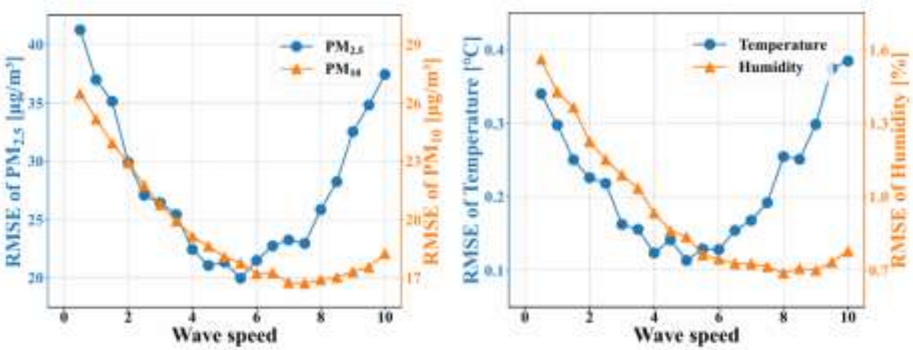
Method	Non-linear	Data-driven	Physics-inspired
KNN [20]	✗	✓	✗
ST-Kriging [19]	✗	✗	✗
GPR [36], [37]	✓	✓	✗
KDE [38]	✓	✓	✗
DMF [17]	✓	✓	✗
IGMC [18]	✓	✓	✗
ImputeFormer [39]	✓	✓	✗
PINO [40]	✓	✓	✓
MCWDD *	✓	✓	✓

4. Experiments

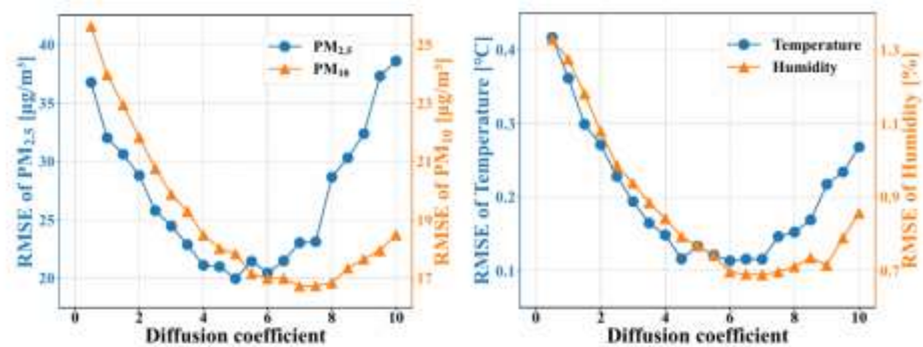
❖ RQ1: How do the hyper-parameters affect the performance of MCWDD

- Impact of Hyper-Parameters:

  wave speed & diffusion coefficient



(a) U-Air (b) Sensor-Scope



(a) U-Air (b) Sensor-Scope

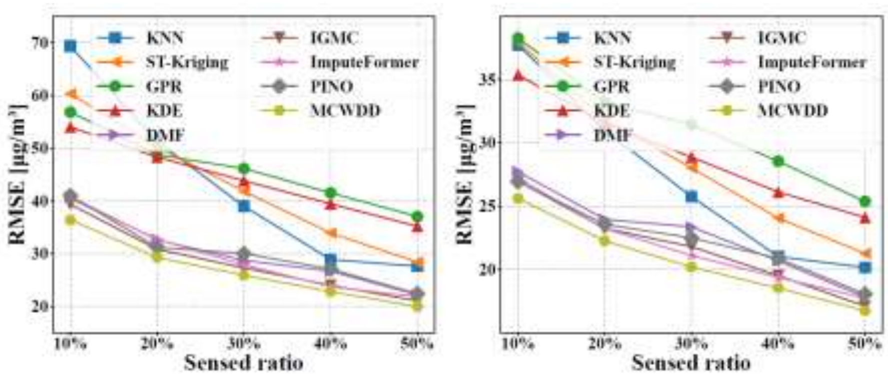
Completion error vs. wave speed

Completion error vs. diffusion coefficient

❖ RQ2: Whether the performance of MCWDD in terms of data completion accuracy exceeds the baseline method

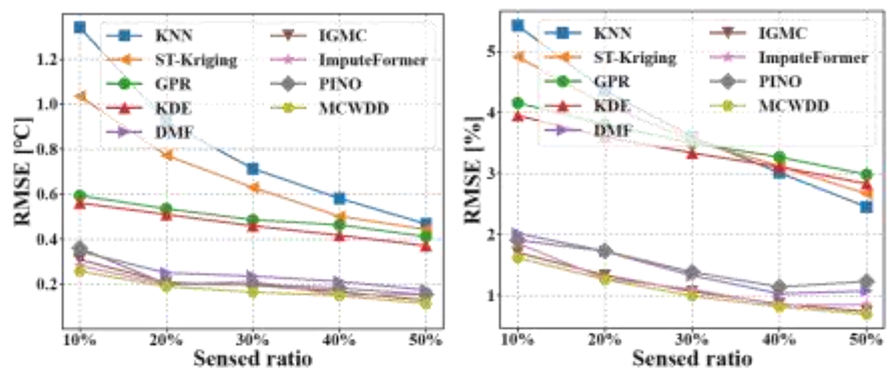
- Robustness Analysis:

Completion error vs. sensed ratios



(a) $PM_{2.5}$ (b) PM_{10}

U-Air

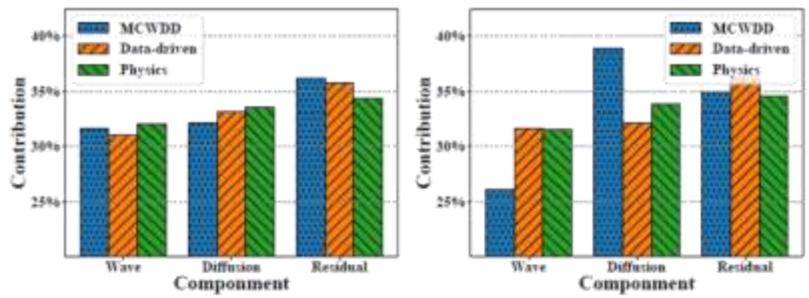


(a) Temperature (b) Humidity

Sensor-Scope

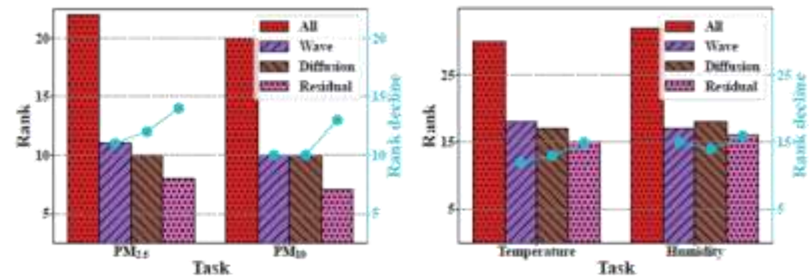
4. Experiments

❖ RQ3: How much impact does each component of WDD have on data completion results



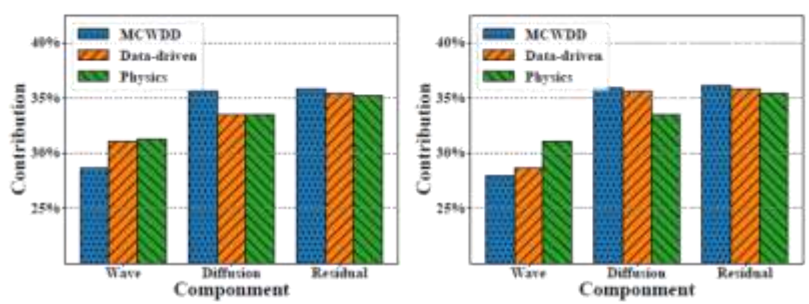
(a) $PM_{2.5}$ (b) PM_{10}

Fig. 9: The impact of data completion contributions from different components over **U-Air**.



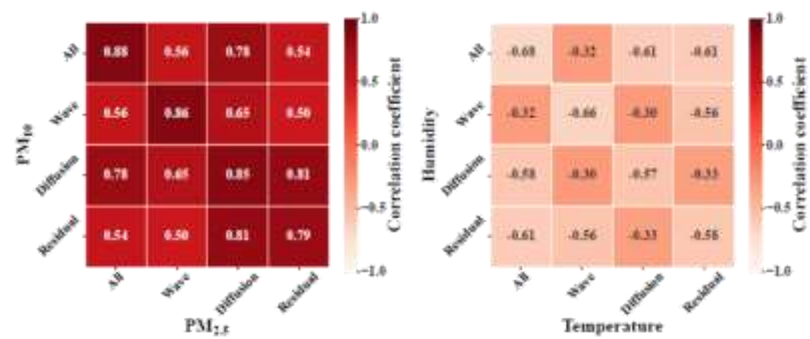
(a) **U-Air** (b) **Sensor-Scope**

Fig. 11: The impact of rank decline by different components over **U-Air** and **Sensor-Scope**.



(a) Temperature (b) Humidity

Fig. 10: The impact of data completion contributions from different components over **Sensor-Scope**.



(a) **U-Air** (b) **Sensor-Scope**

Fig. 12: The correlation coefficient of different components over **U-Air** and **Sensor-Scope**.

❖ RQ4: How does the removal or alteration of specific components affect the overall performance of MCWDD

- Completion error of $PM_{2.5}$/PM_{10} vs. sensed ratios over **U-Air**

Method	Sensed ratio				
	10%	20%	30%	40%	50%
w/o data-driven	39.977	31.037	29.366	27.946	25.705
w/o physics	39.369	30.845	27.311	24.010	21.050
MCWDD	36.399	29.295	25.931	22.807	19.991
Improvement	2.970	1.550	1.380	1.203	1.059

Method	Sensed ratio				
	10%	20%	30%	40%	50%
w/o data-driven	28.469	23.278	22.690	20.589	17.660
w/o physics	26.985	23.246	21.804	19.512	17.105
MCWDD	25.622	22.260	20.199	18.533	16.745
Improvement	1.363	0.986	1.605	0.979	0.360

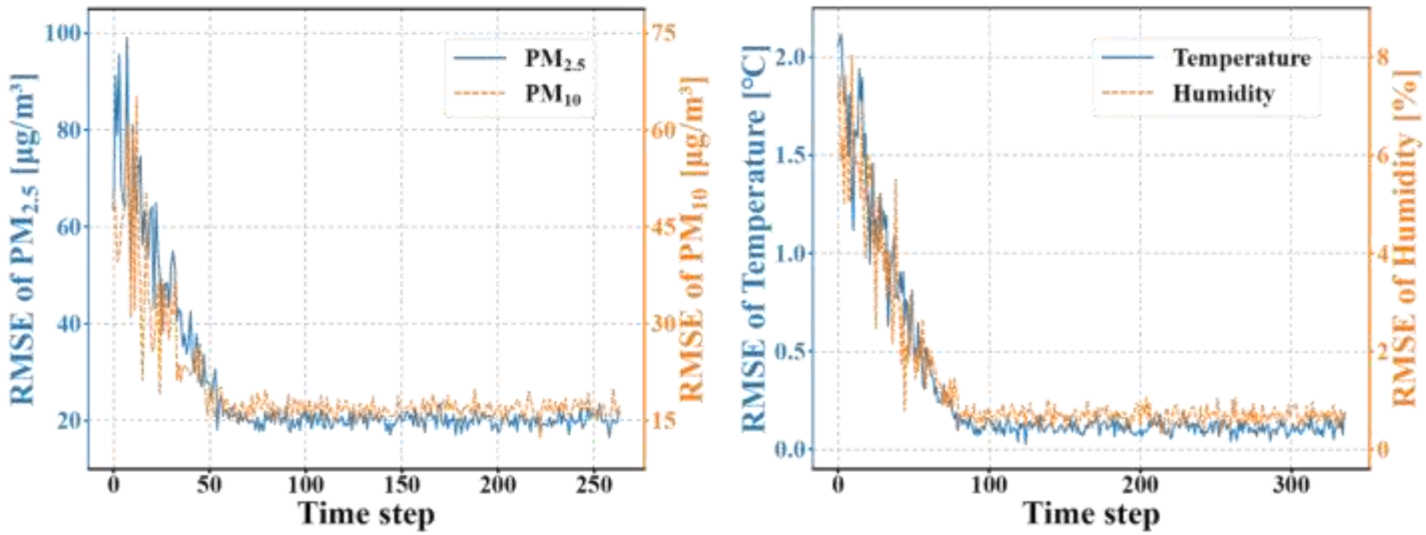
- Completion error of Temperature/Humidity vs. sensed ratios over **Sensor-Scope**

Method	Sensed ratio				
	10%	20%	30%	40%	50%
w/o data-driven	0.3813	0.2327	0.2113	0.1692	0.1302
w/o physics	0.3089	0.2072	0.1958	0.1661	0.1275
MCWDD	0.2573	0.1893	0.1640	0.1489	0.1135
Improvement	0.0516	0.0179	0.0318	0.0172	0.0140

Method	Sensed ratio				
	10%	20%	30%	40%	50%
w/o data-driven	1.8225	1.3484	1.0753	0.9321	0.7709
w/o physics	1.7014	1.3283	1.0613	0.8592	0.7346
MCWDD	1.6161	1.2592	0.9907	0.8133	0.6895
Improvement	0.0853	0.0691	0.0706	0.0459	0.0451

4. Experiments

- RQ5: What is the training cost of MCWDD?
  - Inference accuracy under different input time steps



(a) **U-Air** (b) **Sensor-Scope**

  - Training Time over **U-Air** and **Sensor-Scope**

Dataset / Task	**U-Air**		**Sensor-Scope**	
	$PM_{2.5}$	PM_{10}	Temperature	Humidity
Training Time	107.254	99.369	127.147	117.950

Thank you for listening!

Q&A