AirPhyNet: Harnessing Physics Guided Neural Networks for Air Quality Prediction

Kethmi Hirushini Hettige[1], Jiahao Ji[3], Shili Xiang[2], Cheng Long[1], Gao Cong[1], Jingyuan Wang[3]

[1] School of Computer Science and Engineering, Nanyang Technological University, Singapore

[2] Institute for Infocomm Research, A*STAR, Singapore

[3] School of Computer Science & Engineering, Beihang University, China

ICLR 2024

Air Quality Prediction

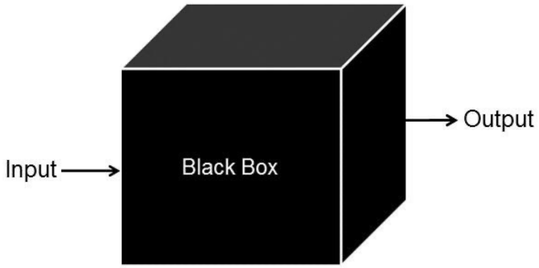
Air Quality is a critical concern which affects human health, environmental sustainability, urban planning, and climate change mitigation. Air quality prediction and modelling plays a pivotal role in public health and environment management.

Complex Air Quality Dynamics

Capturing the influence of multitude of factors such as meteorological conditions.

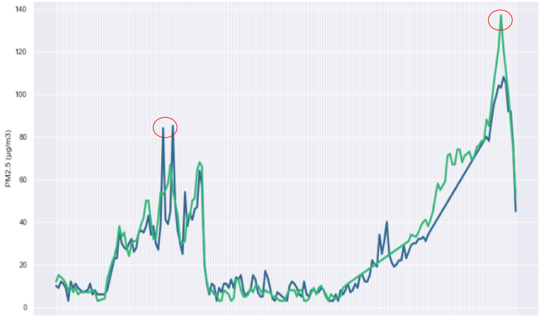
Problem of Data Sparsity

Limited availability of high-quality data, especially in certain regions.



Model Interpretability

Understanding reasoning behind specific predictions is crucial for decision making.



Sudden Change Prediction

Models need to be adaptive and responsive to rapid changes in influential factors.

Related Work

Air Quality Forecasting Methods

Physics Based

Inspired by **atmospheric science** and formulate governing equations that can represent atmospheric processes.

Ex: Chemical Transport Models (CTMs) simulate atmospheric dynamics using ordinary or partial differential equations.

Data Driven

Utilize **historical pollution data** to learn complex relationships.

Ex: Spatiotemporal Graph Neural Networks (STGNN) based networks are prominent while recently attention-based models have emerged as a robust alternative.

Motivation

Limitations of Existing Models

Physics Based	Data Driven
• Involves solving **complex differential equations** and are often employed on larger spatial scales **restricting fine grained predictions**. • Require **parameter calibration** which limit ability to intricate real-world conditions	• Require **extensive training data** to achieve accurate long-term predictions. • Absence of physics constraints **limit their generalizability** • The “black-box” nature of deep learning models raises **challenges on interpretability**

AirPhyNet : A novel hybrid deep learning approach which use underlying physics related to air particle movement as domain knowledge

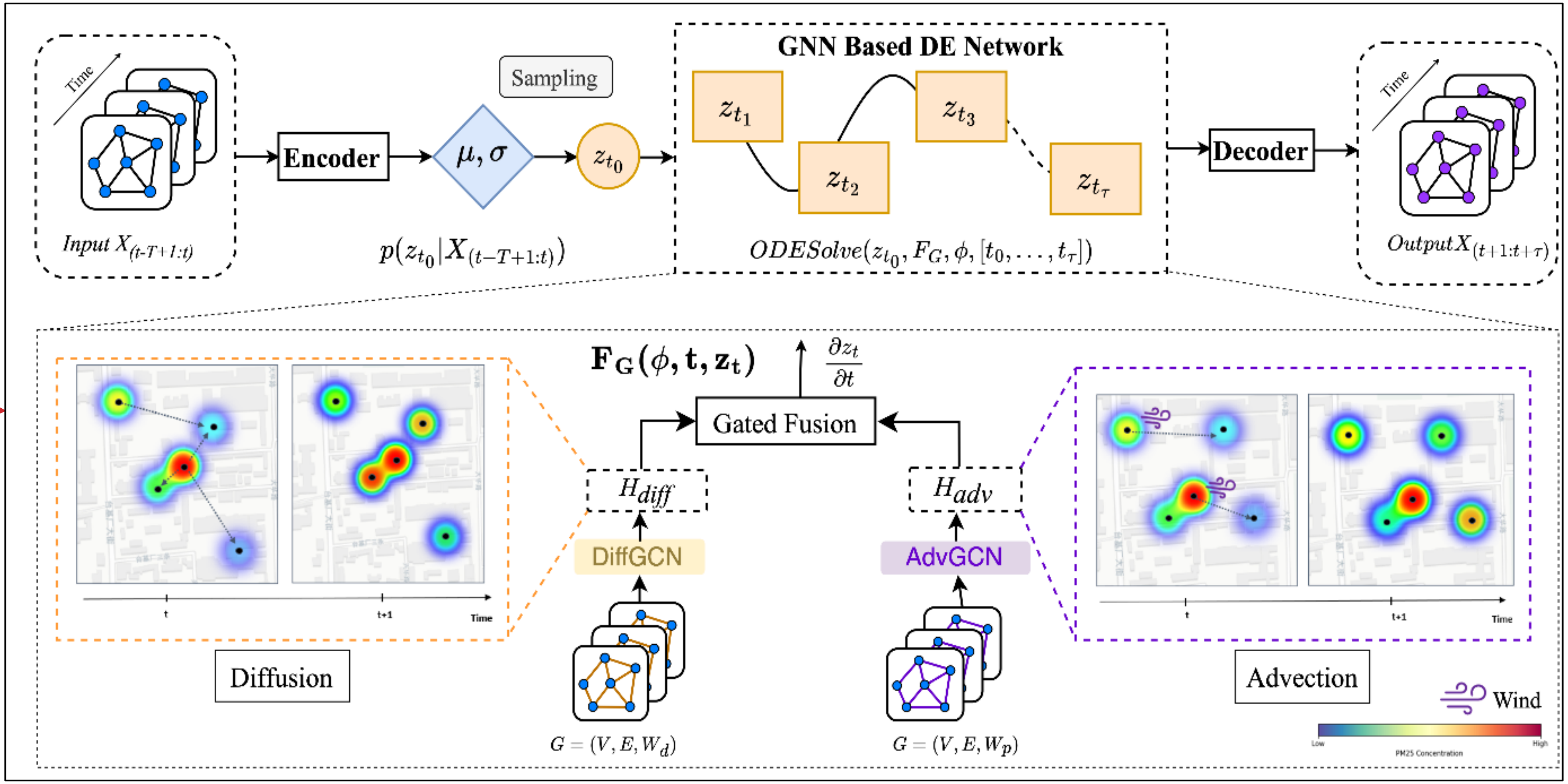
AirPhyNet Model Architecture

RNN-based Encoder :
Encode PM2.5 concentrations into an initial state

GNN based DE network:
Captures the physical dynamics of air pollutant transport

Decoder:

Generates the final PM2.5 concentrations based on the learnt dynamics of physical processes.



Experimental Results : Overall Performance

Model	Beijing Data						Shenzen Data					
	24h		48h		72h		24h		48h		72h	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
HA	38.37	83.91	45.8	95.56	50.58	101.51	8.35	12.52	9.72	14.34	10.54	15.39
VAR	60.10	102.92	60.44	103.02	60.64	103.07	21.50	26.09	21.84	26.50	22.17	26.92
DCRNN	35.99	52.55	49.66	67.50	57.01	74.67	7.76	11.54	9.99	14.11	10.98	15.14
STGCN	33.70	49.16	38.93	54.98	43.93	56.57	7.23	10.61	9.35	12.99	9.97	13.59
GMAN	50.62	66.05	50.73	66.07	50.69	65.87	9.76	12.70	10.08	13.22	10.07	12.99
GTS	34.99	51.45	54.18	71.87	73.50	89.59	6.58	9.55	8.70	12.15	10.54	13.94
PM25GNN	50.94	65.87	48.81	65.64	51.51	66.55	9.90	12.31	10.02	12.63	10.30	12.85
AirFormer	29.62	46.49	38.43	56.52	43.39	58.68	7.24	10.83	9.66	13.47	10.21	13.98
LatentODE	44.83	53.96	45.95	55.44	47.14	57.39	9.85	12.30	10.24	12.89	10.73	13.33
ODE-LSTM	46.19	57.56	49.18	62.39	51.45	63.66	10.55	13.24	11.36	14.18	12.03	14.81
AirPhyNet	**29.11**	**42.16**	**36.69**	**48.66**	**42.23**	**53.07**	**6.38**	**9.48**	**8.18**	**11.38**	**9.51**	**12.51**
% Improvement	1.73	9.33	4.53	11.49	2.66	6.19	3.01	0.76	5.97	6.38	4.58	2.64

- AirPhyNet outperforms all competing baselines in both metrics, across all time horizons for both Beijing and Shenzen data.

- Compared to the second-best method, **on average it achieves reduction in MAE and RMSE by 3.7% and 6.1% respectively**

Experimental Results : Sudden Change Prediction

Sudden changes are defined as instances where PM2.5 concentration exceeds 50 μg/m³ and 20 μg/m³ and exhibits a change of more than ±20 μg/m³ in the following three hours for Beijing Data and Shenzhen Data respectively.

Model	Beijing Data		Shenzhen Data	
	MAE	RMSE	MAE	RMSE
DCRNN	18.616	53.0629	3.3468	8.0377
STGCN	21.821	56.8061	3.2087	8.1773
GMAN	48.2294	80.4201	4.6336	9.9652
GTS	22.5072	63.1401	3.7127	8.1299
PM25GNN	28.196	69.6186	4.9186	10.2459
AirFormer	14.1553	48.4329	3.1317	7.9183
LatentODE	43.5	76.8365	3.549	8.639
ODE-LSTM	38.248	76.2235	3.6311	8.6356
AirPhyNet	**13.4987**	**47.7057**	**2.8904**	**7.3148**
% Improvement	4.64	1.50	7.71	7.62

- AirPhyNet shows the best performance in predicting sudden changes in both datasets.
- Compared to the second-best method, **AirPhyNet shows an average reduction of 6.2% and 4.6% in MAE and RMSE respectively**.

Experimental Results : Sparse Data Prediction

How well our model generalizes to unseen data? We evaluated the models using a smaller training dataset with train:val:test split as 3:1:6.

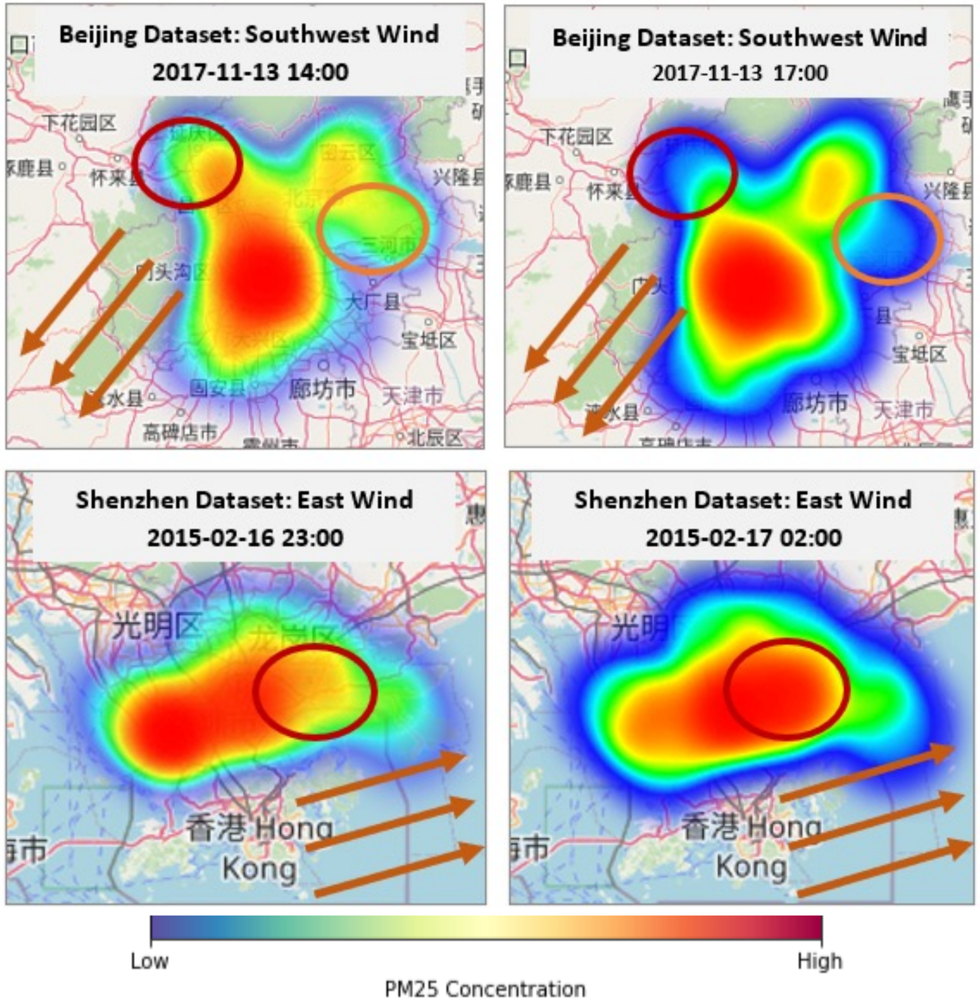
Model	Beijing Data						Shenzen Data					
	24h		48h		72h		24h		48h		72h	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
DCRNN	44.24	56.03	55.80	68.13	64.90	78.24	7.99	11.48	10.11	13.99	11.06	15.01
STGCN	33.73	44.90	39.93	50.51	46.62	57.53	7.25	10.50	8.96	12.13	9.75	12.92
GMAN	46.61	55.59	47.60	56.62	48.65	57.63	9.69	12.35	9.88	12.68	9.91	12.64
GTS	38.09	48.97	60.28	71.45	80.42	92.06	6.60	9.61	8.70	11.87	10.09	13.37
PM25GNN	47.34	61.55	46.33	60.81	47.64	61.65	10.78	13.18	10.84	13.22	10.90	13.23
AirFormer	31.47	43.20	41.64	53.22	44.50	56.39	7.14	10.22	8.93	12.17	9.25	12.51
LatentODE	44.02	51.40	45.94	53.28	47.68	54.51	9.81	12.15	10.01	12.38	10.25	12.54
ODE-LSTM	49.95	61.16	51.93	63.19	53.44	64.19	10.36	13.26	11.01	13.97	11.49	14.39
AirPhyNet	**27.99**	**38.48**	**35.78**	**45.82**	**40.45**	**50.37**	**6.47**	**9.50**	**7.75**	**10.76**	**8.63**	**11.58**
% Improvement	11.08	10.93	10.39	9.29	9.09	10.68	1.88	1.09	10.92	9.34	6.73	7.43

AirPhyNet has outperformed all baselines with reduction in MAE and RMSE by 8.3% and 8.1% on average respectively compared to the second-best method.

AirPhyNet generalizes effectively and achieves precise long term air quality predictions particularly in scenarios with limited data

Case Study – Advection

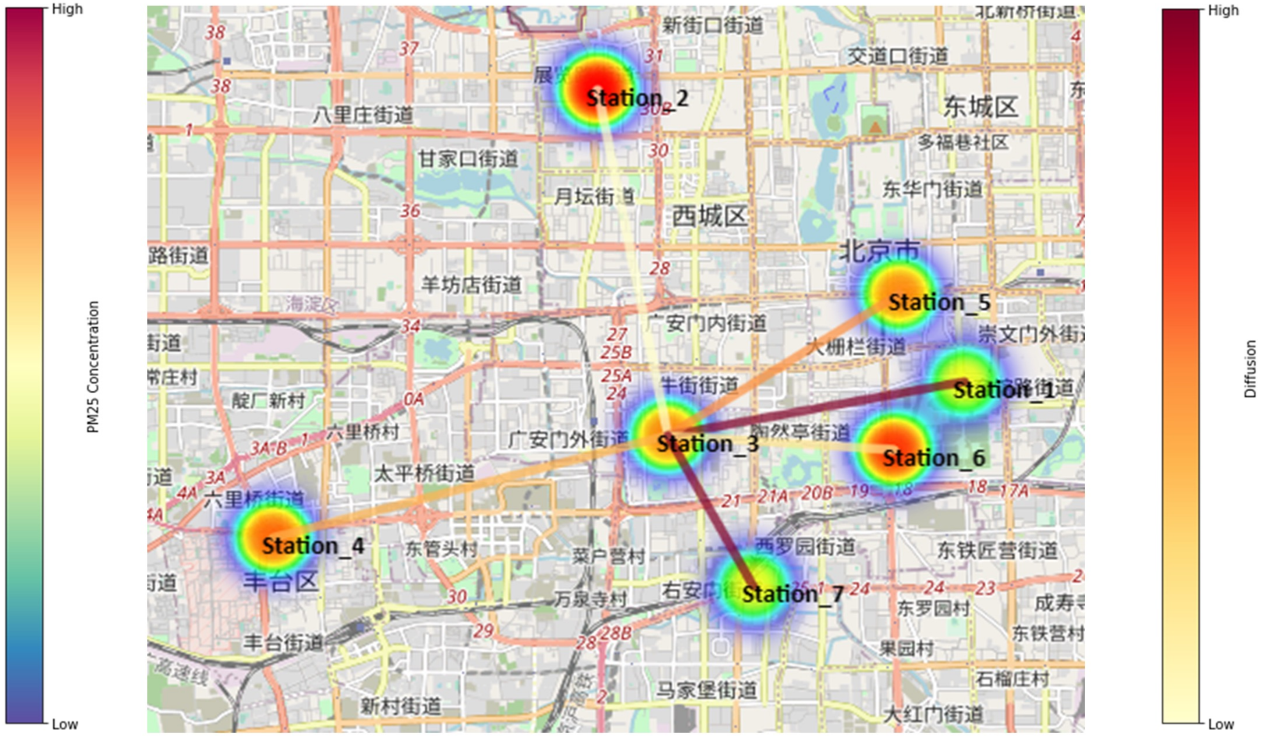
Visualize predicted PM2.5 concentrations and wind direction. Heatmap represents the PM2.5 concentrations and the arrows indicate the wind direction.



- In Beijing dataset, **PM2.5 concentration has reduced** in the circled regions and the particles have joined the **PM2.5 cloud heading towards the southwest wind direction**.

- In Shenzen dataset, **PM2.5 Concentration has increased** over the four-time steps in the circled region and joined the **PM2.5 cloud moving towards the east wind direction**.

Case Study – Diffusion



The figure shows the predicted PM25 concentrations of the AirPhyNet model at particular timestep by a heatmap and the diffusion from station_3 to its neighboring stations by different line segments. The magnitude of diffusion is reflected by the line color.

- The diffusion of particles from Station 3 is relatively higher to stations with low concentrations (Station 1, Station 7) while it's low to stations with higher concentration (Station 2).
- Thus, the predicted concentrations can be interpreted in terms of diffusion.

AirPhyNet successfully captures the underlying physical principles of particle movement and yields precise predictions with a physical meaning.

Conclusion

- Proposal of a novel Physics guided Neural Network known as AirPhyNet, that integrates the fundamental physical processes diffusion and advection into a deep learning framework.

- Experiments conducted on two real world datasets show that AirPhyNet reduced prediction errors by a significant margin (4%-10%) compared to the existing methods in different testing scenarios.

- A case study further affirms model's capability to generate precise predictions that can be interpreted within a physical context.

THANK YOU !