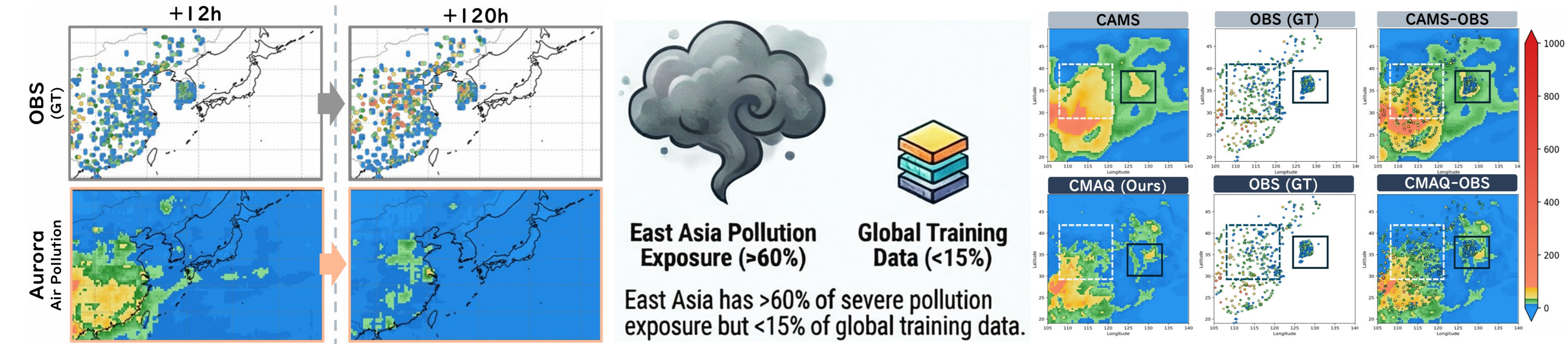


Real-Time Long Horizon Air Quality Forecasting via Group Relative Policy Optimization

Inha Kang¹, Eunki Kim¹, Wonjeong Ryu¹, Jaeyo Shin¹, Seungjun Yu¹, Yoon-Hee Kang², Seongeun Jeong², Eunhye Kim³, Soontae Kim², Hyunjung Shim¹

¹KAIST AI ²Ajou University ³Kunsan University

Problem Thesis: Lack of Localized Dataset & Decision-cost Mismatch



Global foundation models struggle with regional accuracy. It stems from (1) Data used by global models has an error rate 59.5% higher than localized data and (2) Standard objective functions tend to over-predict pollution, leading to frequently and unnecessary alerts.

Experimental Results : Quantitative Results and Qualitative Examples

		PM _{2.5}													
OBS	CMAQ	$\mathcal{L}_{TA}(T=2)$	$\mathcal{L}_{TA}(T=4)$	Overall	+12h	+24h	+36h	+48h	+60h	+72h	+84h	+96h	+108h	+120h	
Aurora Air Pollution [10]					16.06	48.82	40.07	17.75	15.00	4.04	4.01	1.16	1.72	0.46	1.05
✓	✗	✗	✗	50.74	61.53	53.83	51.57	49.87	48.79	48.47	48.15	47.76	47.68	47.65	
✓	✓	✗	✗	54.40	62.56	58.89	56.73	54.42	52.26	51.86	51.36	51.36	51.41	51.38	
✓	✓	✓	✗	57.65	65.89	61.99	60.52	57.58	55.48	54.75	54.04	53.76	53.44	53.34	
✓	✓	✗	✓	59.90	67.53	64.48	61.83	60.39	59.05	58.23	57.44	56.94	56.54	56.21	
				(+9.16)	(+6.00)	(+10.65)	(+10.26)	(+10.52)	(+10.26)	(+9.76)	(+9.29)	(+9.18)	(+8.86)	(+8.56)	

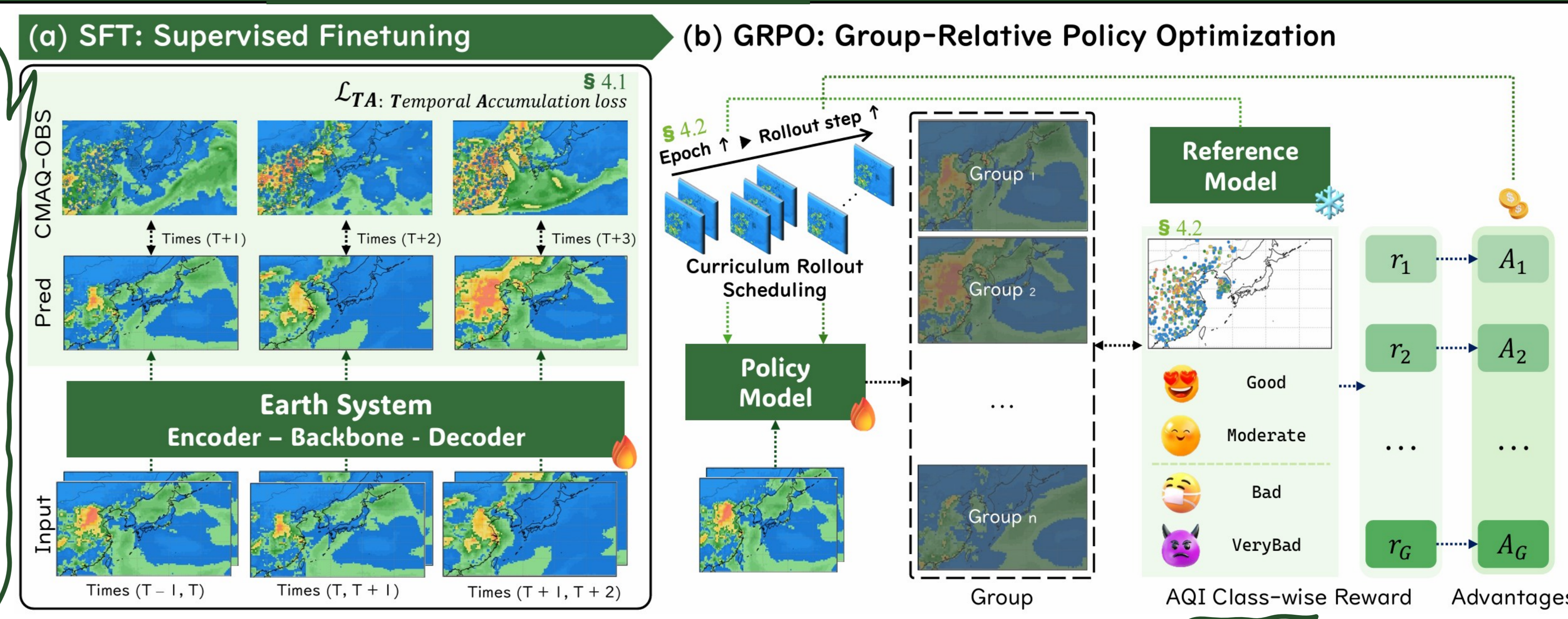
Config	Binary Metrics					AQI (4-class) Metrics										
	Model	Reward	CR	Acc	F1	Prec	FAR	Bias	Acc	F1-macro	F1-weighted	F1-micro	Good	Mod.	Bad	V.Bad
Aurora	-	-	-	68.93	16.06	68.40	2.24	0.13	34.03	23.43	28.91	34.03	42.49	36.82	11.46	2.96
FAKER-Air SFT	-	-	-	69.62	59.90	49.69	32.86	1.52	44.03	41.59	43.63	44.03	38.46	47.79	44.92	35.18
FAKER-Air GRPO	MSE	✗	✗	74.50	48.44	61.90	10.54	0.64	43.73	37.57	42.36	43.73	50.24	46.09	33.93	20.04
FAKER-Air GRPO	AQI	✗	✗	71.40	56.28	52.12	24.19	1.17	42.26	40.75	41.94	42.26	49.19	40.05	39.42	34.35
FAKER-Air GRPO	AQI	✓	✓	74.51	56.72	57.98	17.32	0.96	45.16	41.90	44.66	45.16	50.92	44.90	41.80	29.98

- ✓ **Aurora (Baseline):** Rapidly loses mesoscale spatial structure at long horizons, failing to capture critical high-pollution events.
- ✓ **FAKER-Air SFT:** Successfully restores coherent spatial patterns and transport dynamics across extended horizons, overcoming baseline degradation.
- ✓ **FAKER-Air GRPO:** Prunes false alarm artifacts while strictly preserving high-pollution cores, achieving decision-grade operational reliability.

[Stage 1: Robust SFT]
Motivation: Exposure Bias
 Standard loss causes errors to accumulate over long horizons.
Method: Temporal Accumulation Loss.
 Supervises auto-regressive forecast trajectories rather than single-step snapshots. Penalizes cumulative errors to ensure temporal consistency and stability across long-horizon forecasts.



Methods: FAKER-Air



[Stage 2: Operational Alignment via GRPO]
Motivation: Decision-Cost Mismatch → MSE is cost-symmetric, but real-world risks are asymmetric
Method: Group-Relative Policy Optimization.

- AQI Class-wise Reward:** Applies asymmetric penalties to align predictions with operational priorities
- Group Ranking:** Optimizes the policy by comparing multiple trajectory rollouts and reinforcing the best.
- Curriculum Rollout (CR):** Gradually increases the rollout horizon to stabilize optimization.

