Alibaba Cloud
Worldwide Cloud Services Partner

Alibaba Cloud Intelligence

AI Powered Weather Forecast

Milestones of Alibaba Cloud Intelligence

Alibaba Cloud
Worldwide Cloud Services Partner

2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Establishment	First Application Release : Loan product "sheepdog"	First Internet Customer	Gain ISO27001 Certification	First 5,000 nodes' cluster in China	Go Global	Expediting Globalization SG and US AZ	Intelligence industrial solution Trial in Hangzhou	TOP Partner of IOC	IOT，Edge Computing	Cloud Intelligence Go China	MNC Go Asia	Tokyo 2020 Beijing 2022

Past 10 years

Cloud
- 2009 Few Knew
- 2019 >80% Will

Self-developed Apsara
- 2009 First Cloud OS in China
- 2013 5K Cluster
- 2018 1st Prize of CIE

Intelligence Platform
- 2017 AI Platform
- 2018 More cities in China

Innovation & Cost
- 2009 High cost of IT
- 2019 Cloud-based innovation

Market Share
- 2009 Newly Established
- 2019 No.1 in China No.3 in Global

Global IDC
- 2009 Start from China
- 2022 27 Regions 84 AZ

Core Technology
- 2009 Development of Apsara
- 2020 Apsara/X-Dragon/PolarDB/Flink…

National Affairs
- 2012 First Double 11
- 2015 12306 Travel rush
- 2017 Weibo Peak Flow

Coming 10 years

4 Trend of Digital Infrastructure Modernization

Cloud
Reliable and easy-to-use cloud

Digitization
Big data and intelligence

IoT
Cloud-integrated IoT

Mobile
Mobile collaboration anytime, anywhere

Alibaba Cloud Product Portfolio

Core Internet Technology

Cloud Native
- Container Service for Kubernetes
- Container Registry
- Message Queue for Apache RocketMQ
- Enterprise Distributed Application Service (EDAS)

Video and Traffic
- Apsara Video Live/VOD
- ChatApp
- Edge Compute
- CDN/DCDN
- Short Message Service

IOT
- IoT Cloud Service
- IoT Network Service
- IoT Security Service
- Link IoT Edge
- Device Development Service

Application service
- Domains&Website
- Energy Expert
- Corporate Office Collaboration
- Blockchain as a Service
- eKYC

Security
- Service Security (Managed Security Service)
- Business Security (Game Shield / Content Moderation)
- Data Security (Key Mgmt. Service/Data Encryption Service)
- Infrastructure Security (Anti-DDos/Cloud Firewall/Certificate Mgmt. Service/Security Center/Web Application Firewall)
- Identity Security (ActionTrail/Resource Access Mgmt.)

Data Intelligence

Database

RDS for MySQL	ApsaraDB for Redis	PolarDB for MySQL
RDS for SQL Server	ApsaraDB for MongoDB	PolarDB for PostgreSQL
RDS for PostgreSQL	Data Transmission Service	AnalyticDB for MySQL
RDS for MariaDB TX	Data Management	AnalyticDB for PostgreSQL
DBStack	Database Backup	Data Lake Analytics

Big Data & AI

MaxCompute	E-MapReduce
DataWorks	Realtime Compute for Apache Flink
Elasticsearch	Quick BI
Hologres	PAI
Vision AI	Image Search

Cloud Infrastructures

Elastic Compute
- Elastic Compute Service
- Elastic Bare Metal Instance
- Simple Application Server
- Elastic GPU Service
- Elastic Desktop Service
- Dedicated Host

HPC
- Super Computing Cluster
- Elastic High Performance Computing
- Batch Compute

Storage
- Object Storage Service
- Block Storage
- File Storage
- Apsara File Storage NAS
- Log Service
- Tablestore

Network

Cloud Data Transfer	Cloud Enterprise Network
Express Connect	Smart Access Gateway
NAT Gateway	Global Acceleration
Server Load Balancer	Elastic IP Address
VPN Gateway	Shared Bandwidth

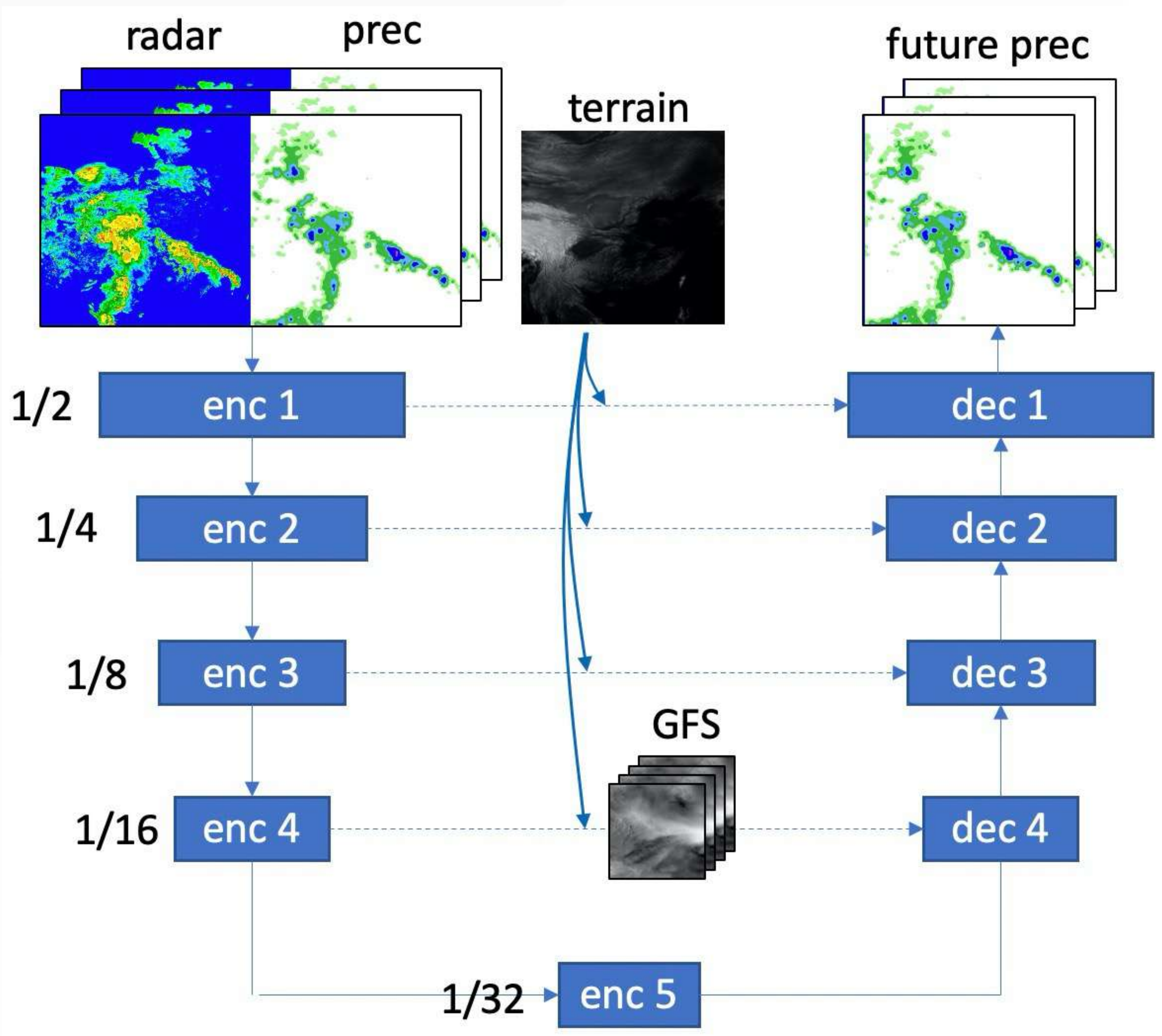
Hybrid Cloud
- Apsara Stack
- Cloud Box
- Local Region
- SOFAStack

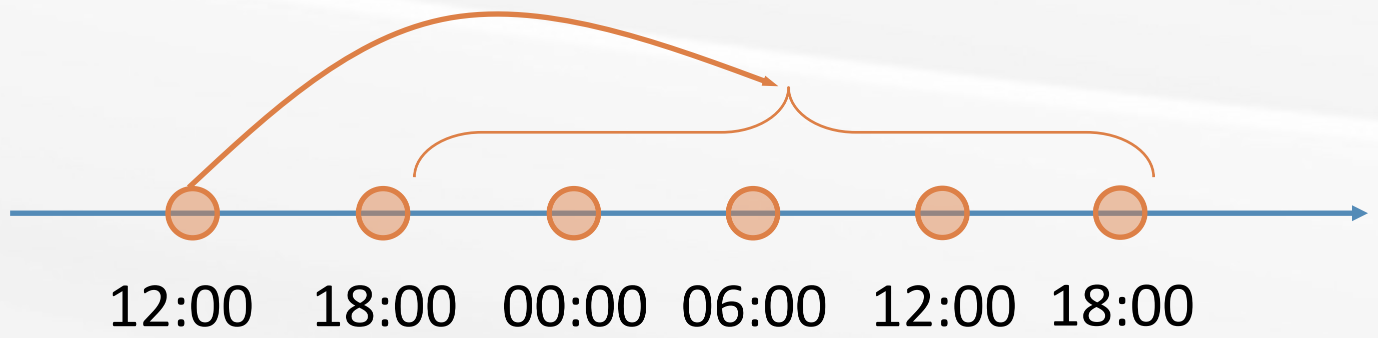
6 hours DL-based Nowcasting: longer and more accurate

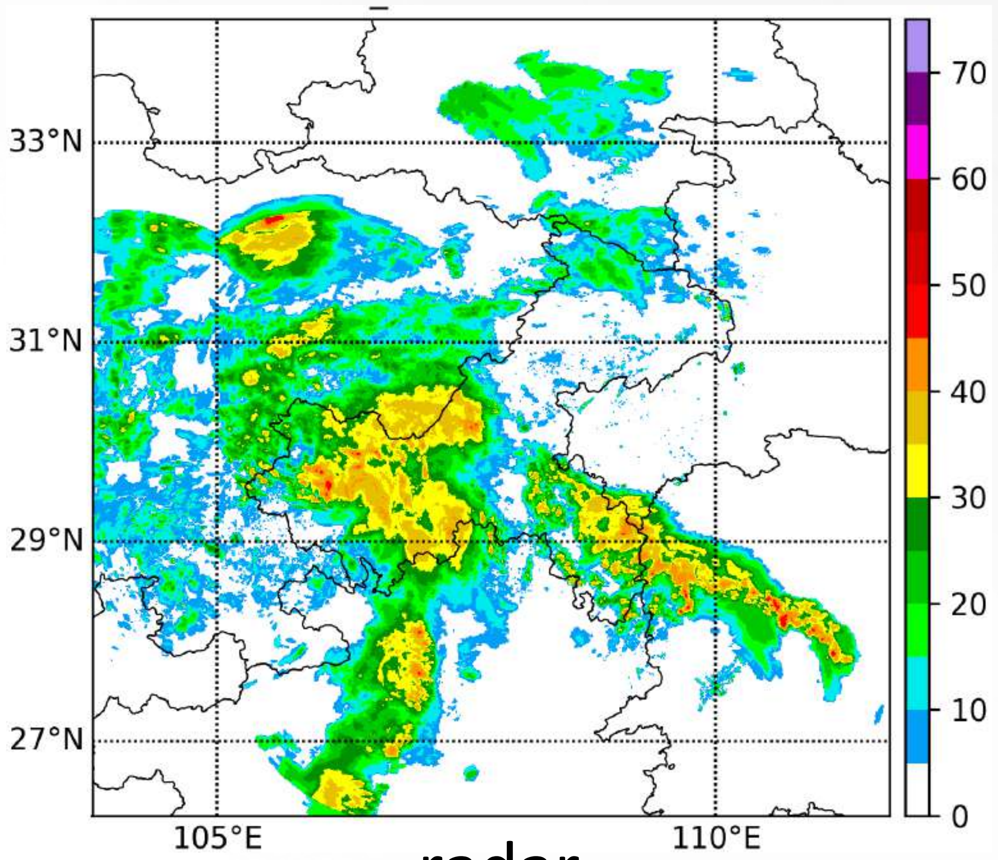
Alibaba Cloud
Worldwide Cloud Services Partner

Multi-source data

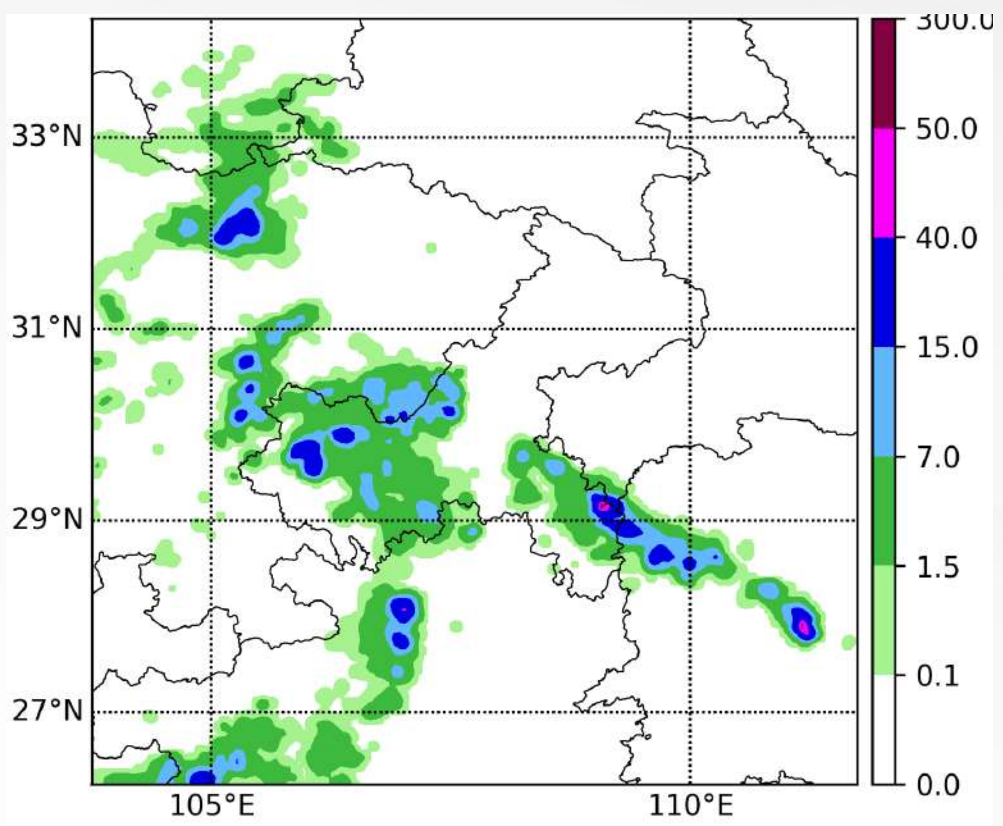
Data	Spatial resolution	Temporal resolution
Composite radar reflectivity	1km	10min
Precipitation	5km	10min
NCEP GFS	25km	1h
Topography	1km	–



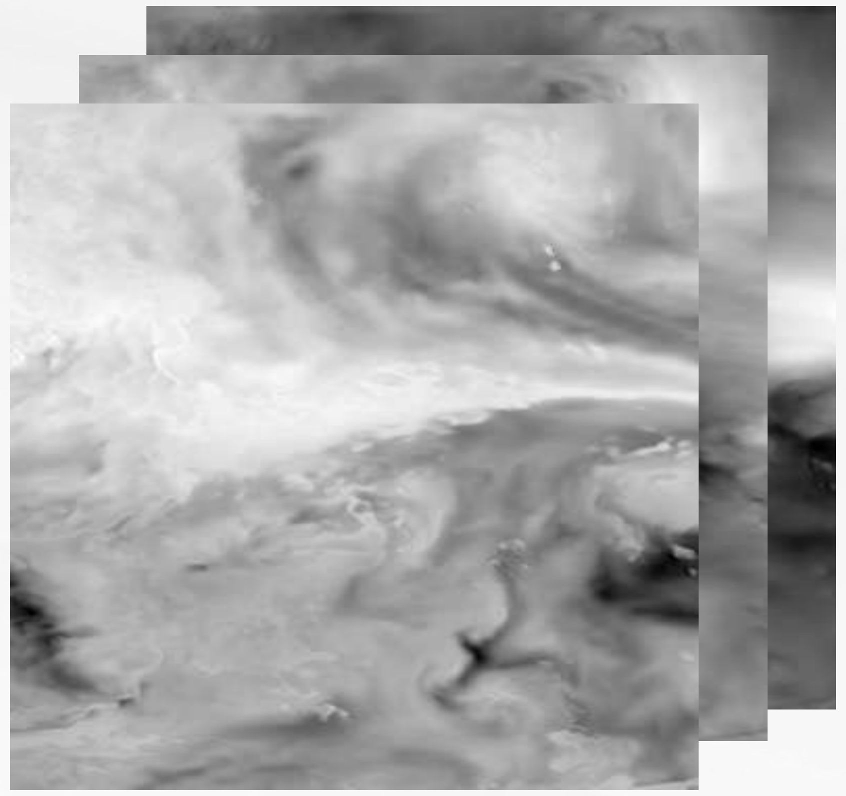




radar



precipitation



GFS (u/v/q)

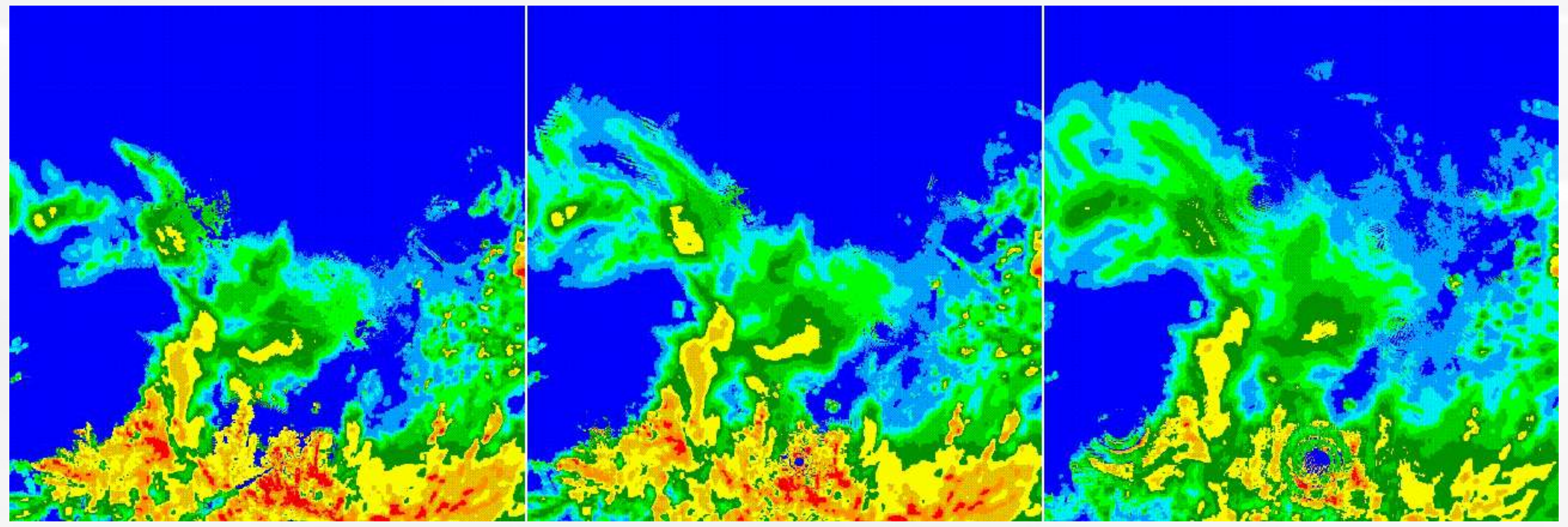
Model	GFS	Lead time	TS_0.1	TS_5	TS_10	TS_20	TS_30
UNet		3h	0.4551	**0.2896**	**0.2194**	0.1670	0.1260
		6h	0.3264	0.1973	0.1548	0.1055	0.0748
	✓	3h	**0.4969**	0.2869	0.2098	**0.1789**	**0.1448**
	✓	6h	**0.3979**	**0.2147**	**0.1665**	**0.1498**	**0.1147**

Deployed at National Meteorological Center of CMA

Nowcasting using radar base reflectivity of different levels

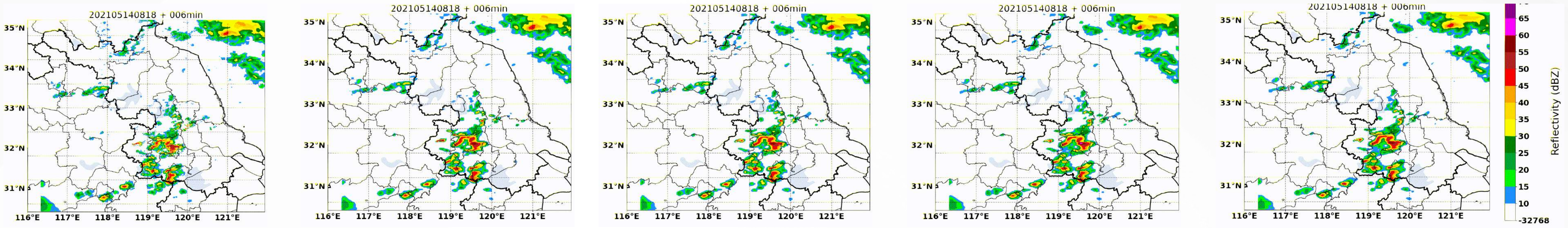
Conclusions

- Multi-level input is better than single level: data from convection development level facilitate the prediction of storm strength and trend
- Multi-level output is better than single level: more information to be supervised
- Channel attention further boost forecast performance



0.5km | 2.5km | 6km

Model		Input levels	Target level	Eval. level	Ch. Atten.	# Params	Lead time	TS20	TS30	TS40	BIAS20	BIAS30	BIAS40
UNet	A	2.5 km	2.5 km	**2.5 km**		30.45 M	3h	0.4099	0.2654	0.1244	0.8064	0.5746	0.5914
	B	(0.5, 2.5, 6) km	2.5 km	**2.5 km**		30.45 M	3h	**0.4202**	**0.2749**	**0.1356**	0.7889	0.5771	0.7234
	C	(0.5, 2.5, 6) km	(0.5, 2.5, 6) km	**2.5 km**		30.53 M	3h	**0.4260**	**0.2799**	**0.1410**	0.7636	0.5582	0.7011
	D	(0.5, 2.5, 6) km	(0.5, 2.5, 6) km	**2.5 km**	✓	31.46 M	3h	**0.4447**	**0.3051**	**0.1519**	**0.8625**	**0.6564**	**0.9100**



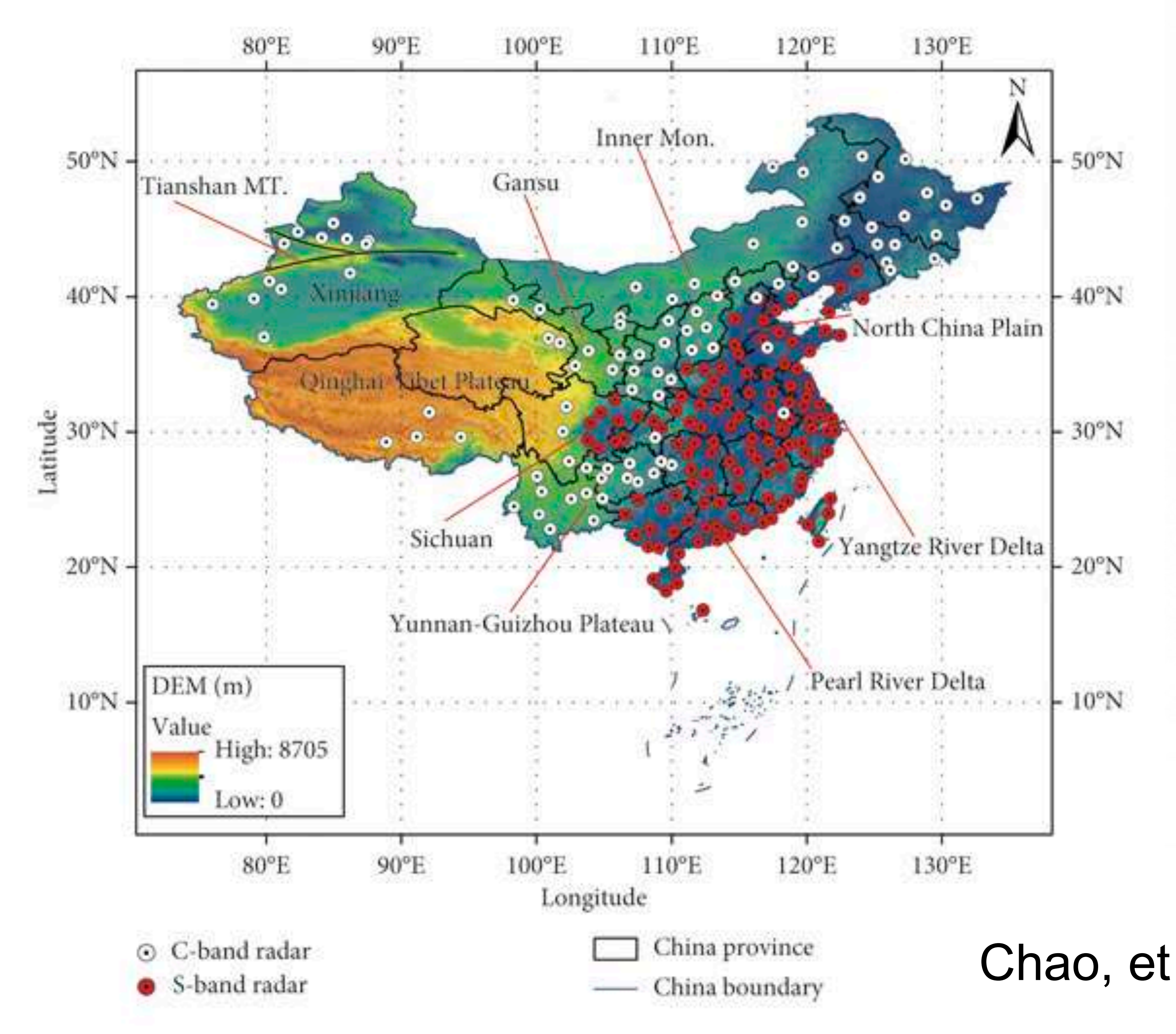
ground truth | A (2.5km) | B (2.5km) | C (2.5km) | D (2.5km)

Preliminary Results on Using Satellite Data for Nowcasting

Motivation:

- Lack of observing and monitoring infrastructure, including radar in LDCs and SIDS
- Terrain blockage of radars in mountainous area
- Satellite data has better spatial coverage than radar data

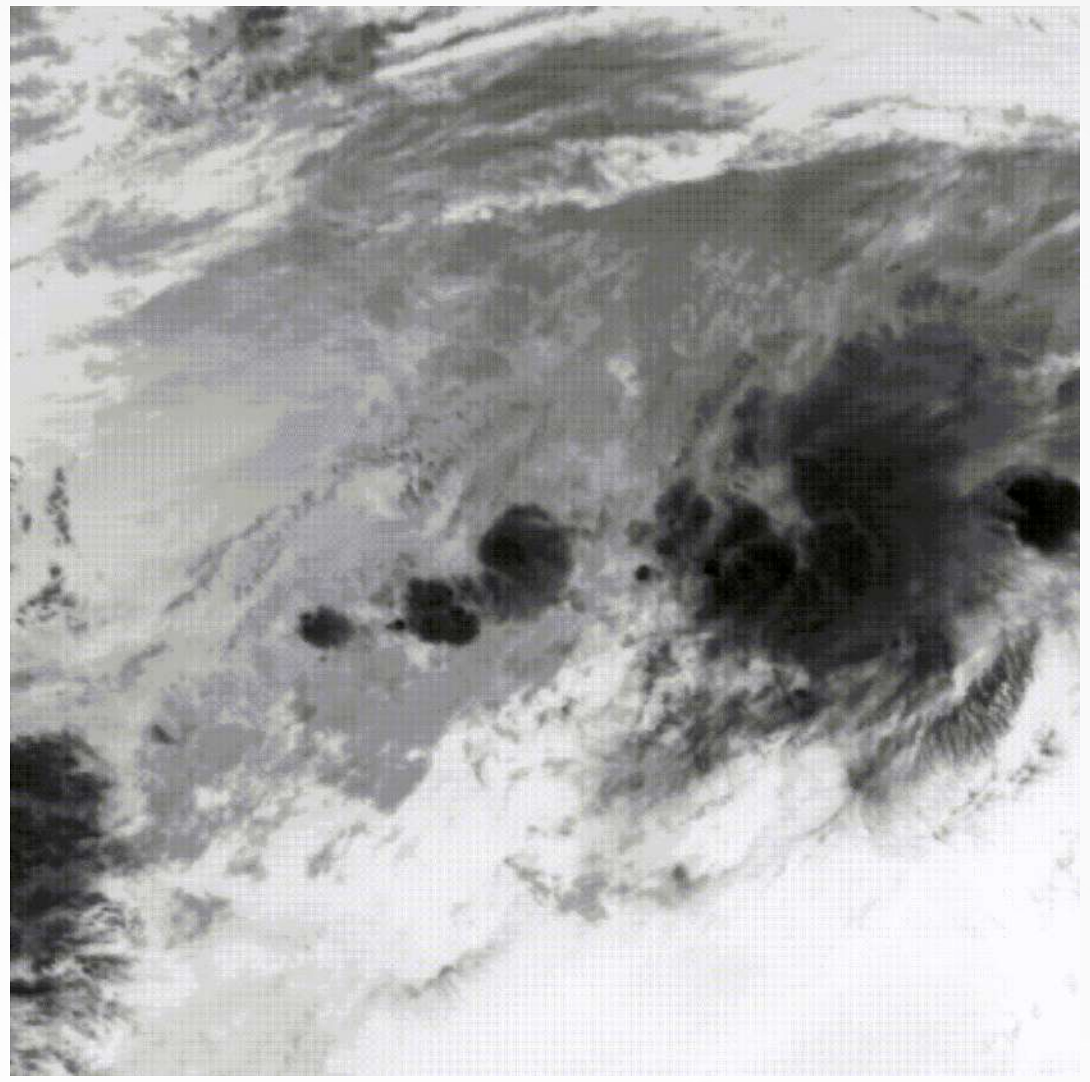
Sparse radar distributions in the western China



Chao, et al 2019

Promising results using satellite data for nowcasting

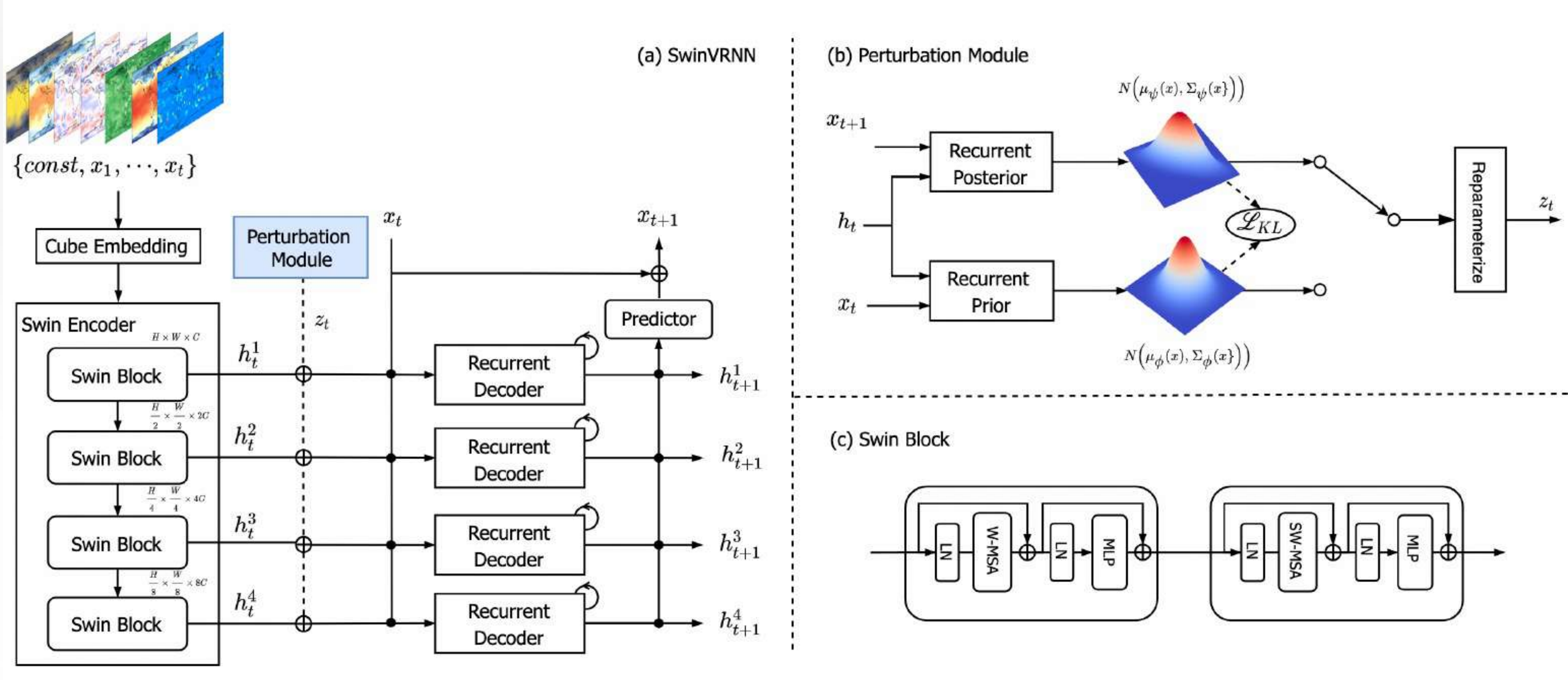
NO	Radar	Satellite	TS20	TS35	TS45	TS50
A	☑		0.366032	0.174292	0.044409	0.027988
B		☑	0.300933	0.140818	0.043290	0.026523



Pure Data-Driven Weather Forecasting Model: SwinVRNN

- SwinVRNN Model = Deterministic SwinRNN + Perturbation Module

Overall architecture of SwinVRNN



- Difference of SwinRNN from previous convolutional RNNs (e.g., ConvLSTM, ConvGRU):

1) Recurrent design: extracts features from all the historical frames at once, more memory-efficient
2) Memory update: Replace gates with self-attention
3) Multi-scale prediction: Use features of different spatial resolutions
4) Residual prediction: forecast 6 hour changes in variables instead of variables themselves

- Data: 6 hourly ERA5 reanalysis data of 5.625º from WeatherBench dataset (Rasp et al., 2020)

- Training and testing:

1979 – 2016 for training, 2017 – 2018 for testing

36 hour historical input training on *5 days and test up to 14 days*

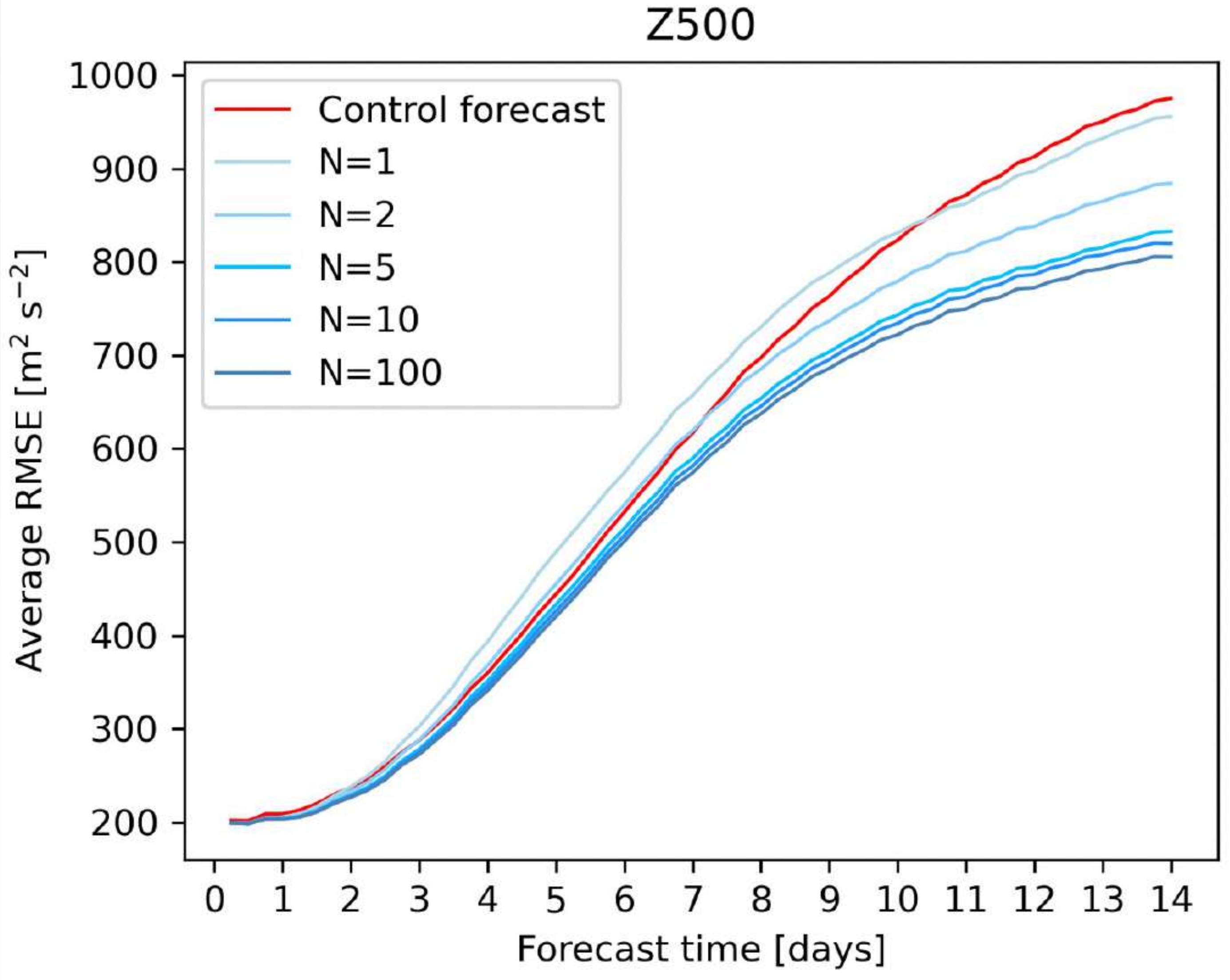
			RMSE (5 days)			
Residual Add	Multi-scale	Swin Attention	Z500 (m^2s^{-2})	T850 (K)	T2M (K)	TP (mm)
✓	✓	✓	428	2.20	1.75	2.17
✓		✓	431 (+3)	2.22 (+0.02)	1.80 (+0.05)	2.21 (+0.04)
	✓	✓	450 (+22)	2.29 (+0.09)	1.88 (+0.13)	2.21 (+0.04)
✓	✓		489 (+61)	2.41 (+0.21)	1.88 (+0.13)	2.18 (+0.01)

Note. Latitude-weighted RMSE scores for 5 day forecasts of Z500, T850, T2M, and TP are reported. The lower the better.

Pure Data-Driven Ensemble Weather Forecasting Model: SwinVRNN

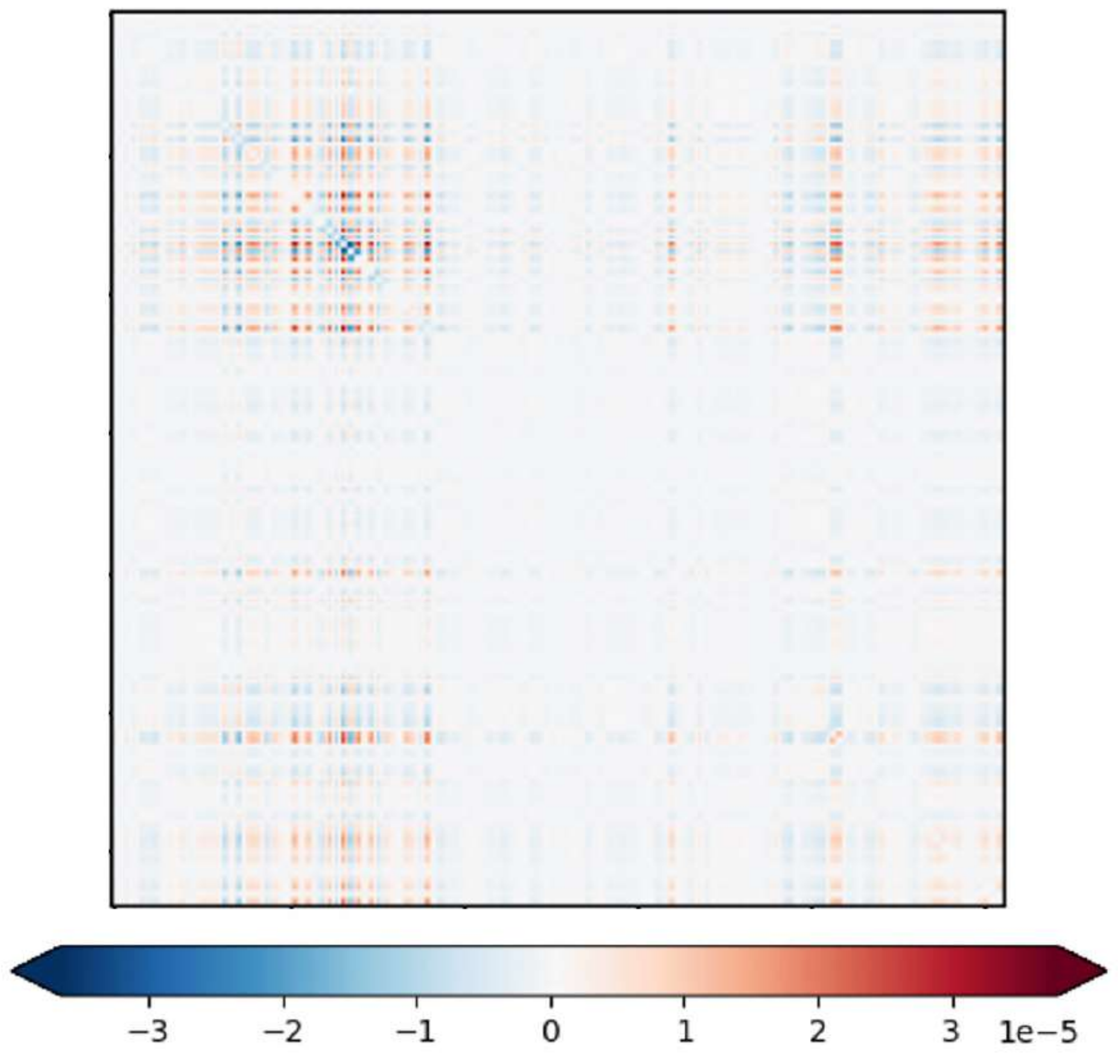
Perturbation Methods:

1) Fixed distribution perturbation (FourCastNet)
2) **Learned distribution perturbation**: spatially and variable dependent
3) Model parameters perturbation: apply MC dropout in both training and inference to mimic the perturbed physics ensembles
4) Multi model ensemble: tune the architecture of SwinVRNN, e.g. depth of the Swin decoder.



(a)

Example of a learned covariance matrix



(b)

Advantage:

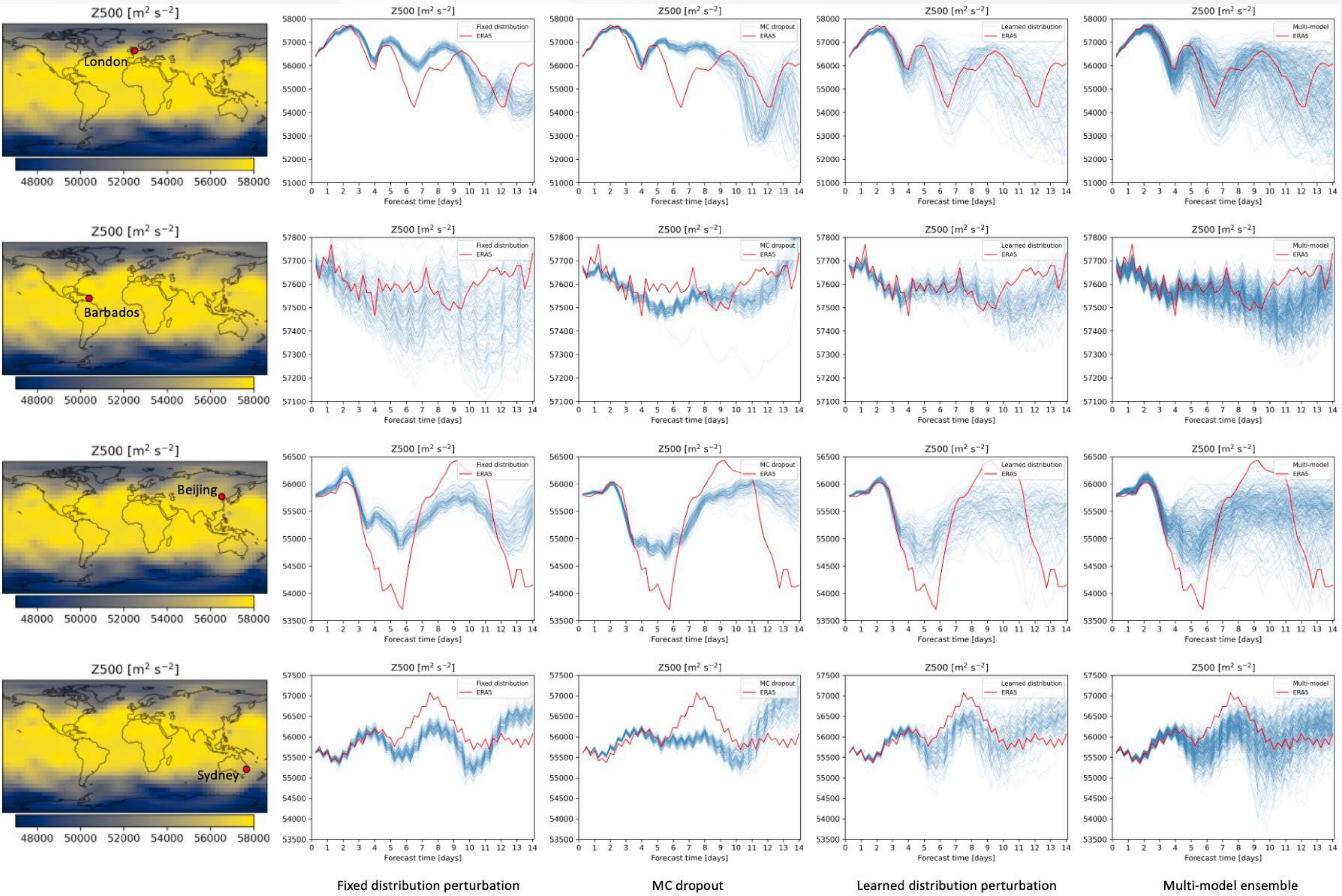
- **Speed**: 44 seconds for 100 member forecasting using a single NVIDIA Tesla V100 GPU

		RMSE (5 days/14 days)				CRPS (5 days/14 days)			
Backbone	Perturbation Method	Z500 (m^2s^{-2})	T850 (K)	T2M (K)	TP (mm)	Z500 (m^2s^{-2})	T850 (K)	T2M (K)	TP (mm)
SwinRNN	-	431/972	2.20/4.25	1.75/3.28	2.17/2.40	339/764	1.76/3.44	1.40/2.63	0.52/0.66
SwinRNN	fixed distribution	431/952	2.20/4.13	1.75/3.25	2.17/2.36	303/646	1.66/2.94	1.27/2.28	0.50/0.58
SwinRNN	MC dropout	426/909	2.18/3.94	1.73/3.08	2.18/2.32	306/597	1.59/2.72	1.25/2.25	0.49/0.52
SwinVRNN	learned distribution	409/803	2.10/3.51	1.68/2.73	2.16/2.31	242/483	1.31/2.18	1.05/1.67	0.45/0.49
SwinVRNN	multi-model	**401/795**	**2.07/3.47**	**1.64/2.70**	**2.16/2.30**	**232/473**	**1.26/2.14**	**0.99/1.63**	**0.44/0.48**
Clare et al. (2021)	MC dropout	627/-	2.91/-	-	-	1500/-	1.69/-	-	-
IFS ensemble	-	297/-	1.73/-	1.57/-	2.15/-	127/-	0.83/-	0.70/-	0.47/-

Note. Latitude-weighted RMSE and CRPS for 5 day and 14 day forecasts of Z500, T850, T2M, and TP are reported. The lower the better.

Pure Data-Driven Ensemble Weather Forecasting Model: SwinVRNN

Fixed distribution | MC dropout | Learned distribution | Multi-model



Ensemble spread for Z500 prediction over London, Barbados, Beijing and Sydney.

Advantage:

- **Accuracy:** outperform IFS on T_{2M} and TP variables up to 5 days at spatial resolution of 5.625º

	RMSE (3 days/ 5 days/ 14 days)			
Method	Z500 (m^2s^{-2})	T850 (K)	T2M (K)	TP (mm)
Weekly climatology	816	3.50	3.19	2.32
T42	489/743/–	3.09/3.83/–	3.21/3.69/–	
T63	268/463/–	1.85/2.52/–	2.04/2.44/–	
IFS	**154/334**/–	**1.36**/2.03/–	1.35/1.77/–	2.36/2.59/–
Naïve CNN	626/757/–	2.87/3.37/–		
Cubed UNet	373/611/–	1.98/2.87/–		
ResNet (pretrained)	284/499/–	1.72/2.41/–	1.48/1.92/–	2.23/2.33/–
FourCastNet	240/480/–	1.50/2.50/–	1.50/2.00/–	2.20/2.50/–
SwinRNN	207/392/882	1.39/2.05/3.90	**1.18**/1.63/2.98	**2.01/2.14**/2.34
SwinVRNN	219/397/788	1.47/2.06/3.45	1.25/1.66/2.69	2.06/2.17/**2.31**
SwinVRNN*	211/388/783	1.43/**2.02/3.43**	1.21/**1.61/2.67**	2.05/2.16/**2.31**

Comparison with SOTA

Pure Data-Driven High-Resolution Weather Forecasting Model: SwinRDM

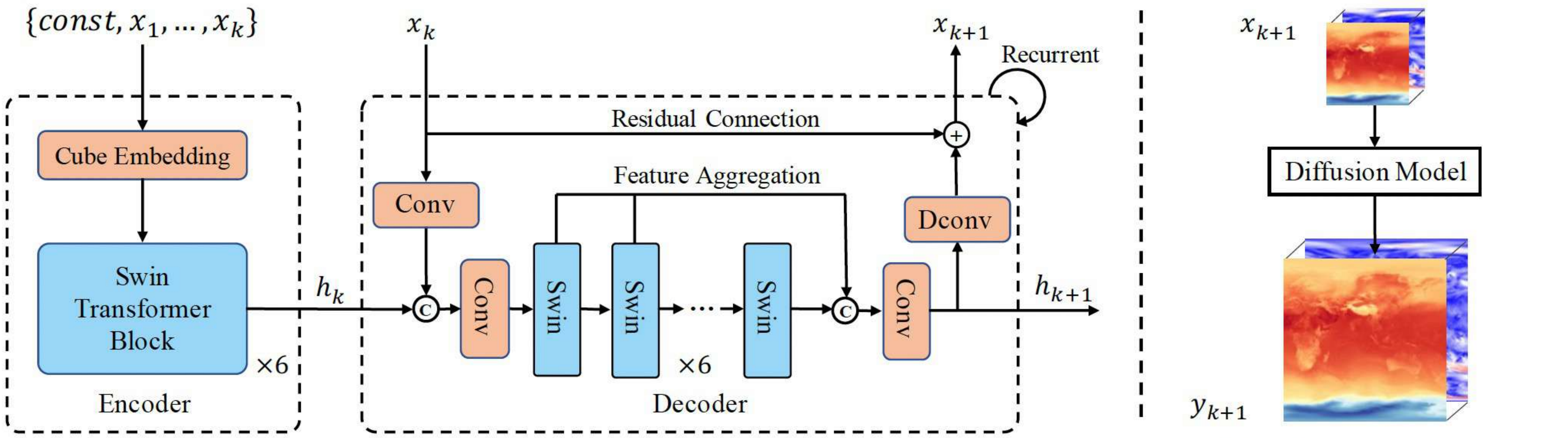
SwinRDM = SwinRNN+ and Conditional diffusion model

✓ SwinRNN+: Improved version of SwinRNN:

- Replace multi-scale network with single-scale network with higher feature dimension
- Multi-Layer Feature Aggregation

✓ Diffusion super-resolution model

Improve resolution from 1.40525º to 0.25º



(a) The Architecture of SwinRNN+ (b) Diffusion-based Super-Resolution

Dataset: 6 hourly ECMWF ERA5 data at spatial resolution of 0.25º, 1979 – 2016 for training, 2017-2018 for testing

Feat. Dim.	Feat. Fusion	Z500	T850	T2M	TP
128		456	2.354	2.196	2.265
128	✓	386	2.050	1.926	2.182
256		394	2.092	1.957	2.207
256	✓	371	1.971	1.843	2.148

Effectiveness of multi-layer feature aggregation

Feat. Dim.	Multi-Scale	Mem.	Params	Z500	T850
128		12.0G	4.0M	417	2.182
128	✓	13.1G	62.5M	386	2.050
256		14.1G	15.5M	374	1.998
256	✓	23.7G	248.6M	371	1.971
384		19.4G	34.8M	359	1.929
512		25.2G	61.6M	354	1.912

multi-scale and single-scale comparison

Super-resolution model comparison

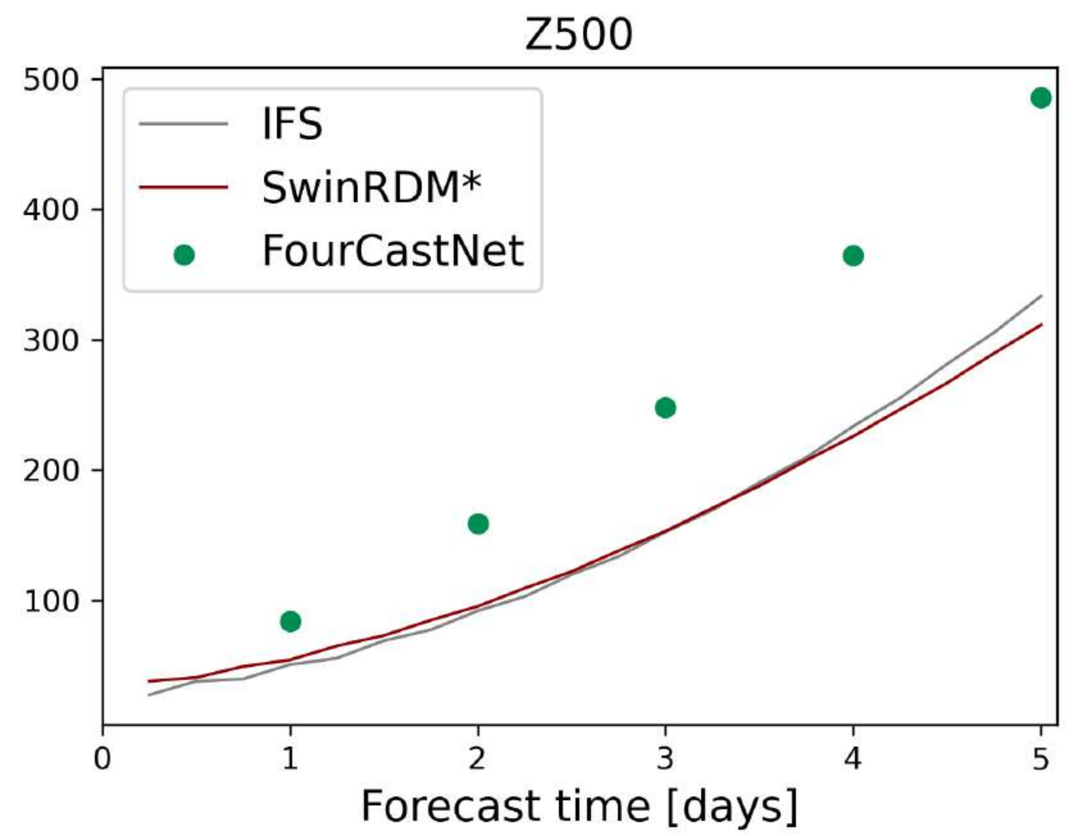
Methods	CSI2	CSI5	CSI10	CSI20	CSI50
Bilinear	0.171	0.051	0.006	0.000	0.000
SwinIR	0.186	0.069	0.013	0.001	0.000
SwinRDM	0.246	0.179	0.111	0.046	0.010
SwinRDM*	0.262	0.190	0.123	0.049	0.006

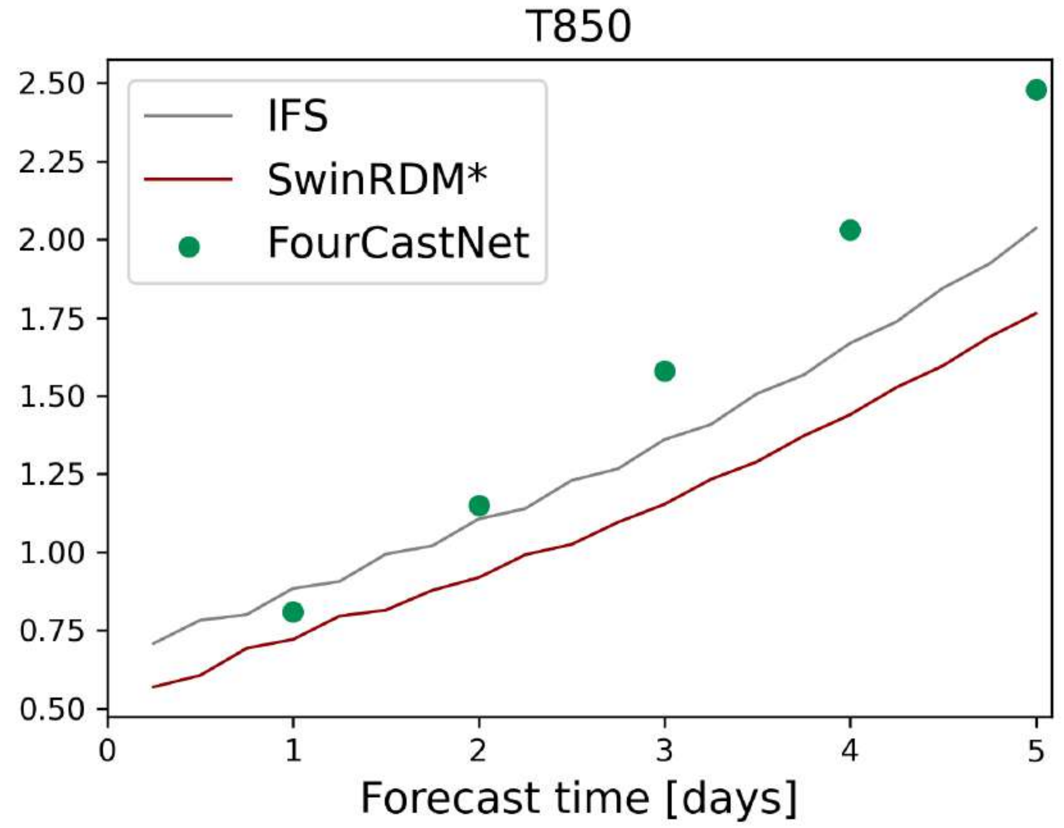
Pure Data-Driven High-Resolution Weather Forecasting Model: SwinRDM

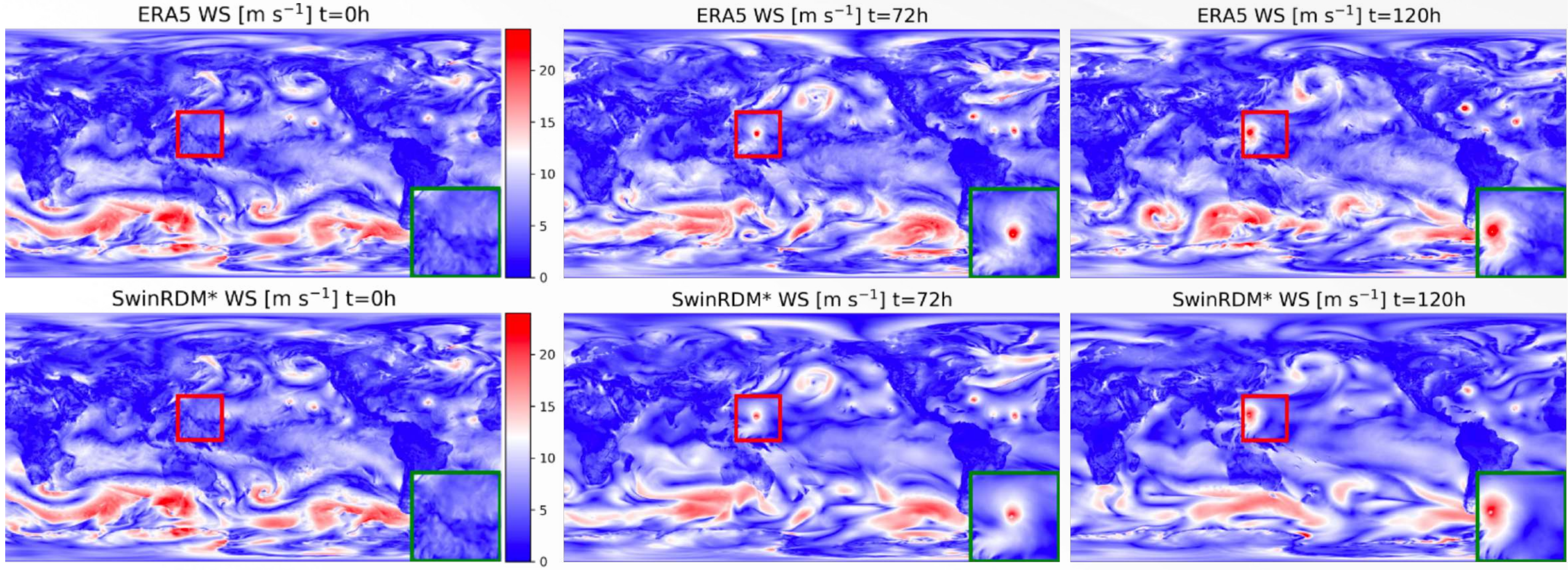
Conclusions:

- SwinRDM outperforms ECMWF IFS at the spatial resolution of 0.25 for 500 hPa geopotential, 850 hPa temperature, 2-m temperature, and total precipitation, at lead times up to 5 days.

Methods	Z500	T850	T2M	TP
IFS	154/334	1.36/2.03	1.35/1.77	2.36/2.59
SwinRNN	207/392	1.39/2.05	1.18/1.63	2.01/2.14
SwinRNN+	152/316	1.12/1.75	0.99/1.42	1.88/2.07
SwinRDM	156/316	1.23/1.83	1.07/1.49	2.02/2.24
SwinRDM*	**153/313**	**1.15/1.76**	**1.01/1.43**	**1.87/2.06**







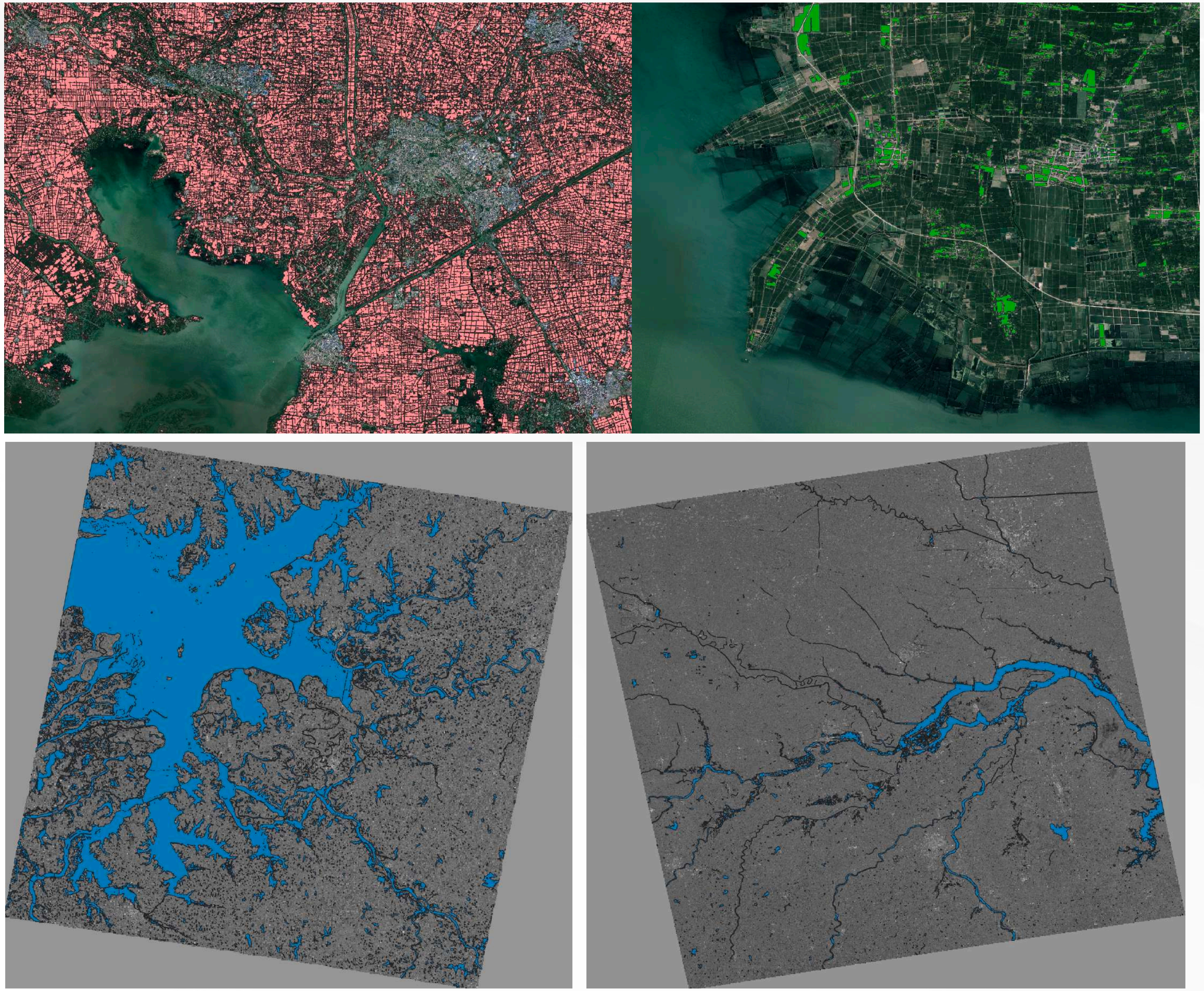
SwinRDM is able to predict Super Typhoon Mangkhut

DL-based Water Bodies Recognition using Synthetic-aperture radar (SAR)

Alibaba Cloud
Worldwide Cloud Services Partner

Motivation

It is too time consuming for human experts to track the changes in size of water body, more efficient methods are required, especially during flood period.



What we did for the Ministry of Water Resource of China in 2020

- Use SAR over satellite images: Works 24 hours and all weather conditions such as cloud, fog, rain, or snow, etc.

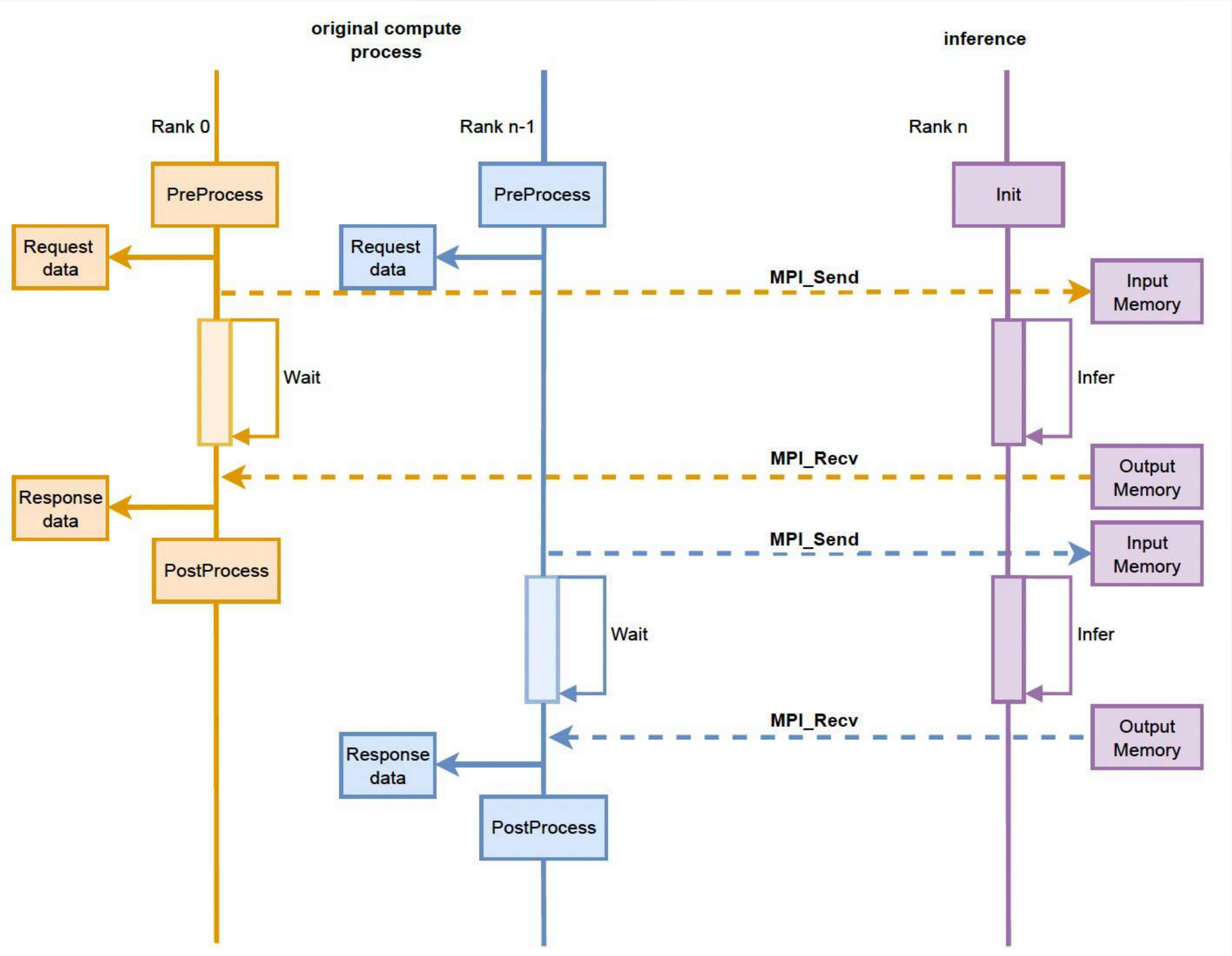
- Transfer learning applied to make the model available within 1 week

WRF-DL: A Bridge between WRF and Deep Learning Parameterizations

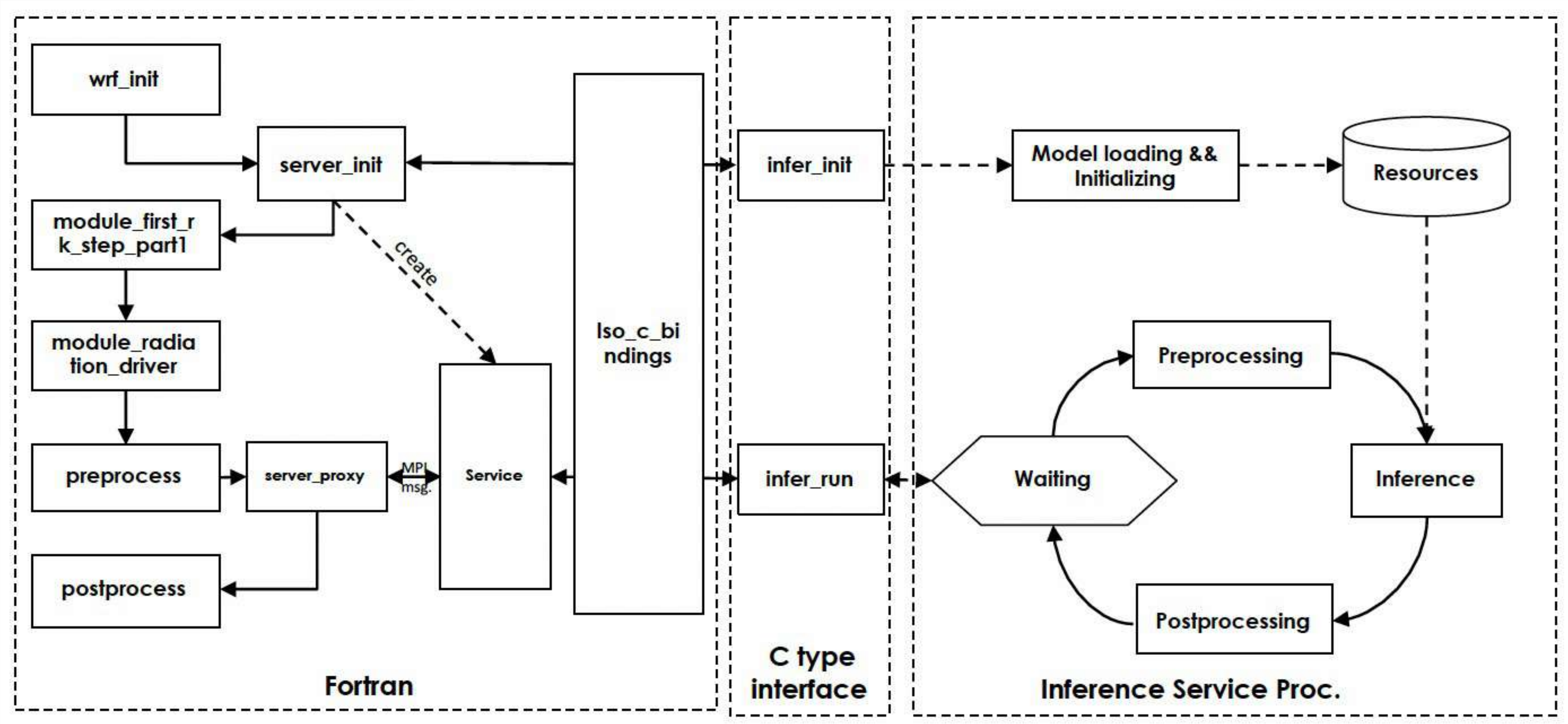
Methodology:

- Use the C Foreign Function Interface (CFFI) to call Python function in Fortran
- Similar to WRF I/O quilt processors, DL model inference can be done using either with exclusive processors or the same processors for WRF calculations

Synchronous sequence implementation between WRF and DL-based parameterization



Framework of WRF-DL coupler



Alibaba Cloud
Worldwide Cloud Services Partner

Alibaba Cloud Intelligence

Tech for Innovation