臺灣區域再分析資料：發展與規畫

連國淵、林怡萱、沈彥志、趙子瑩

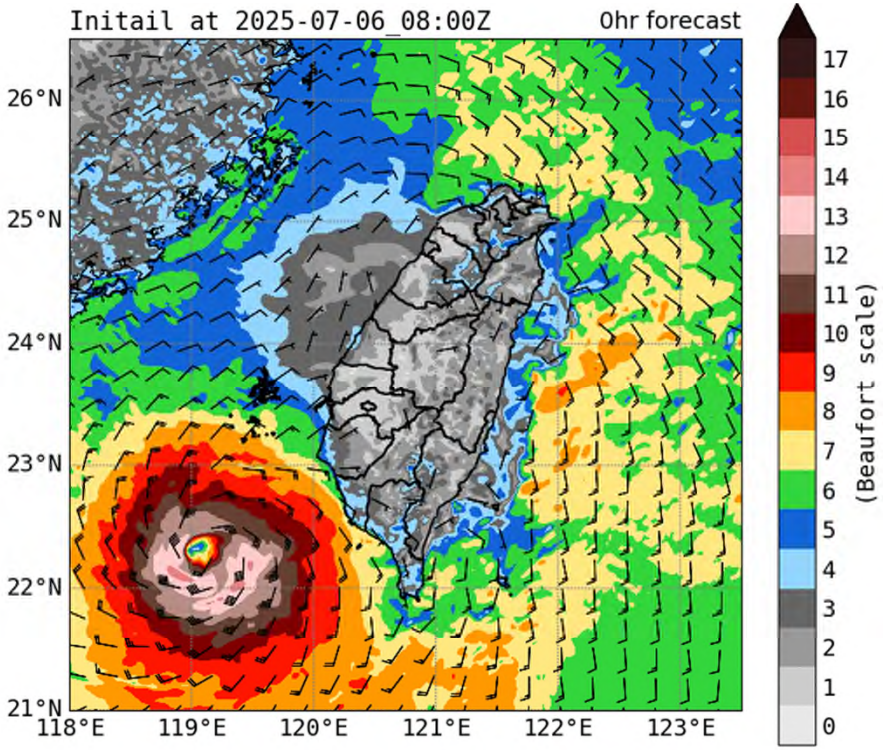
中央氣象署科技發展組

2025/09/02
114 年天氣分析與預報研討會

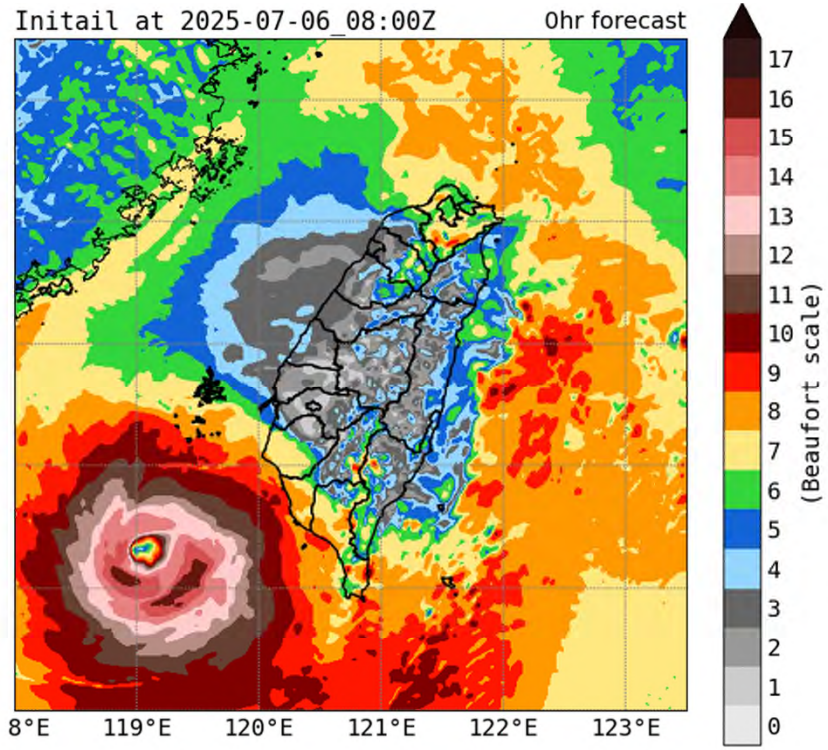
Sample of Taiwan Regional Reanalysis (not final version)

– Typhoon Danas (2025)

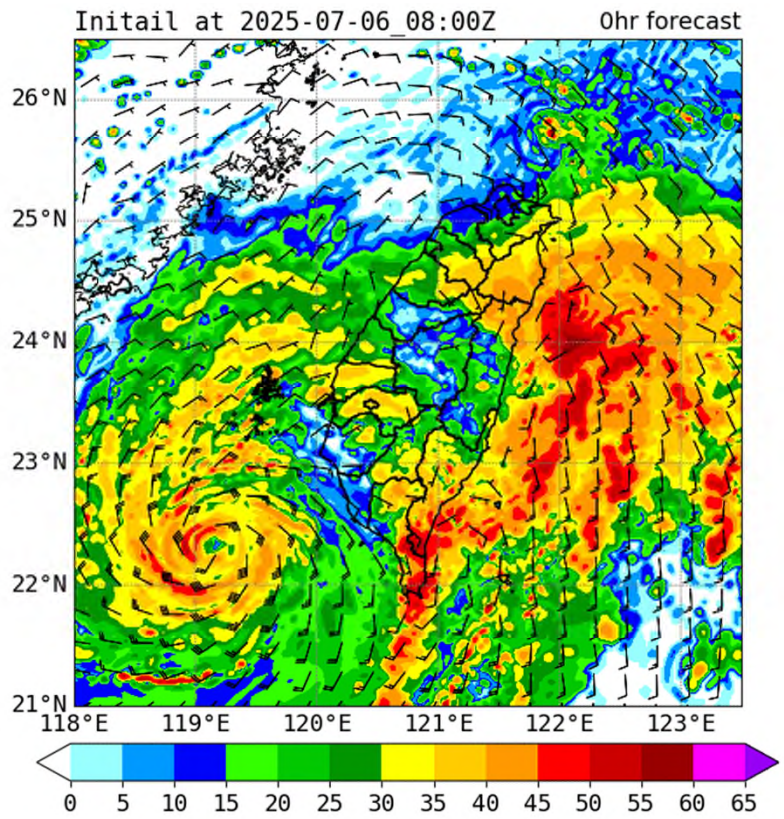
10-m wind speed (Beaufort scale)



Surface wind gust (Beaufort scale)



Simulated max. radar reflectivity (dBZ)



(Atmospheric) Reanalysis

- Using a system **similar to the operational NWP** (model + data assimilation) to perform **retrospective runs,** aiming at generating **analysis fields for past periods.**

- **Purpose:**

  1. To obtain analysis data for historical periods that are **of higher quality than the operational analyses** at the time, by employing a more advanced NWP system nowadays.

  2. To obtain long-term analysis data that are **of relatively consistent quality** using a fixed-version NWP system.

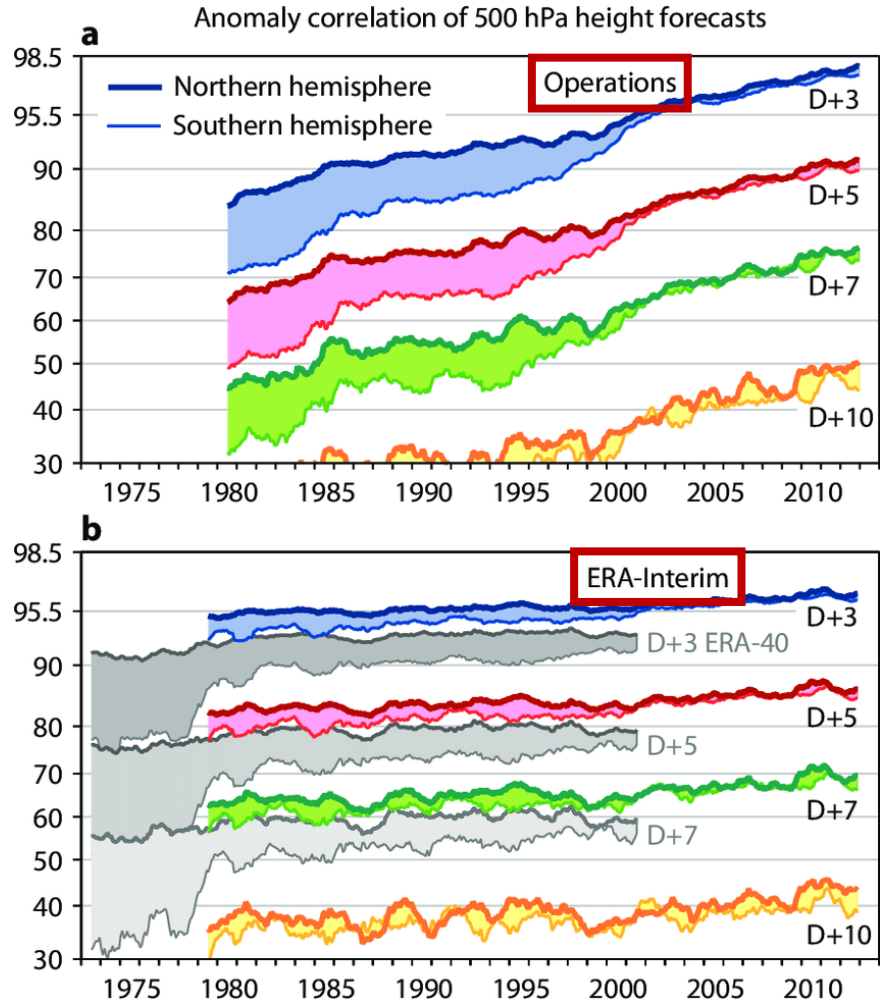
     ***Note**: The quality of the observing systems inevitably varies over time.*

Reanalysis datasets (global and regional)

- **Global reanalysis data (international) :**
  - **NCEP**: R1, R2, CFSR
  - **ECMWF**: ERA-40, ERA-interim, ERA5
  - **JMA**: JRA-55
- **Regional reanalysis data (international) :**
  - **EU**: Copernicus Regional Reanalysis for Europe (CERRA)
    Europe area; 5.5-km grid
  - **DWD**: COSMO Regional Reanalysis (COSMO-REA)
    German area; 2-km grid
  - Japanese “ClimCORE” project
- **Domestic regional “reanalysis-like” data:**
  - **TCCIP’s Taiwan ReAnalysis Downscaling data (TReAD)**
    - WRF dynamical downscaling based on ERA5
    - Taiwan area; 2-km grid; 40+ years

ECMWF Reanalysis

	Operational analysis	Reanalysis		
		ERA-Interim	ERA5	ERA6 (planned)
IFS version	49r1 (2024)	31r2 (2006)	41r2 (2016)	49r2 (2024)
Model resolution	**9 km**	**79 km**	**31 km**	**14 km**
Data assimilation method	Hybrid 4DVar (50 members)	4DVar	Hybrid 4DVar (10 members)	

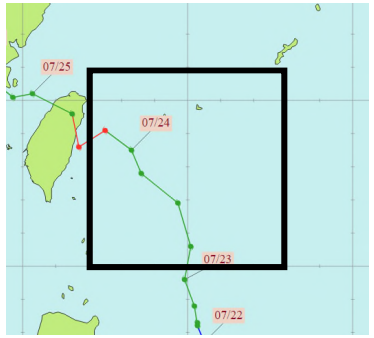


(Dee et al. 2011, ERA-interim)

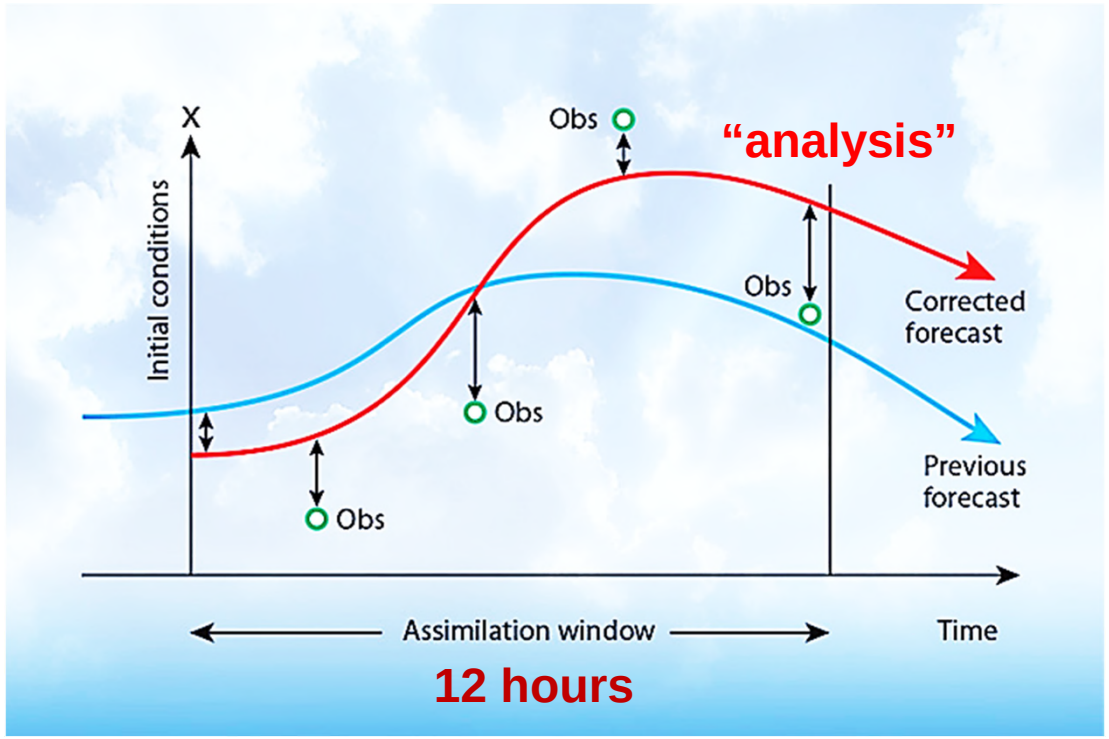
Training datasets for MLWP models

ML model	Training dataset	Data length	Spatial resolution	Temporal resolution
FourCastNet (NVIDIA)	ERA5	~ 40 years	0.25°	6 h
Pangu-Weather (Huawei)				24 h, 6 h, 3 h, 1 h
GraphCast (Google)				6 h
FengWu				
FuXi				
AIFS (ECMWF)				
Aurora (Microsoft)	ERA5, IFS, IFS-ENS, GFS, GEFS, CMIP6, MERRA-2, CAMS	> 100 years in total	0.25°, 0.1°	
NeuralGCM (Google)	ERA5	~ 40 years	0.7°, 1.4°, 2.8°	
GenCast (Google)			0.25°	12 h

Temporal smoothness issue in ERA5

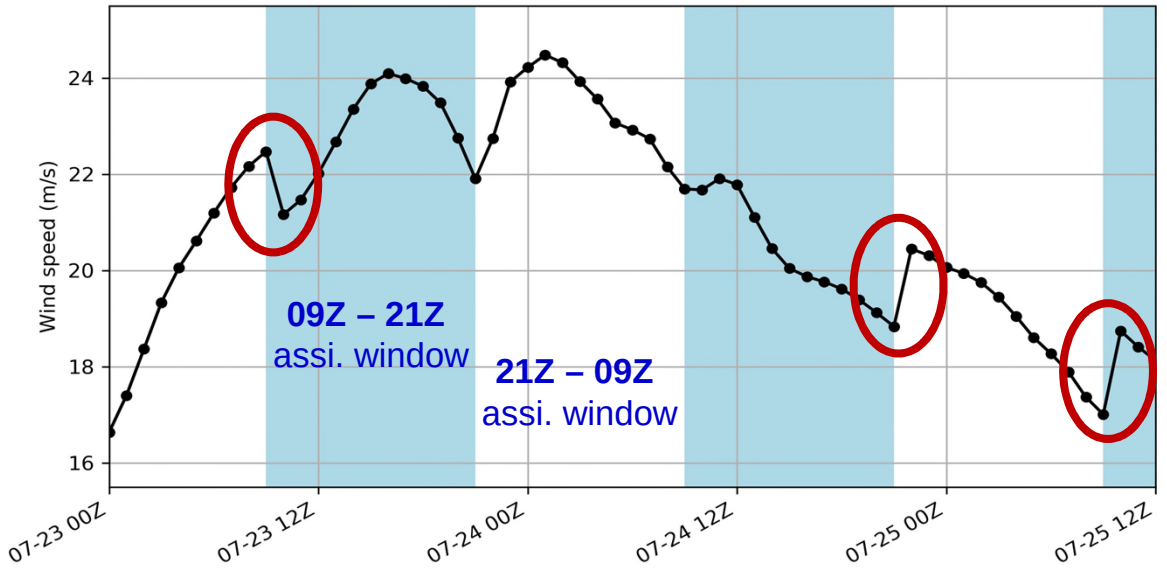


4DVar in ECMWF



https://www.ecmwf.int/en/about/media-centre/news/2022/25-years-4d-var-how-machine-learning-can-improve-use-observations

Example:
Time series of area-averaged wind speed
in the vicinity of Typhoon Gaemi (2024)

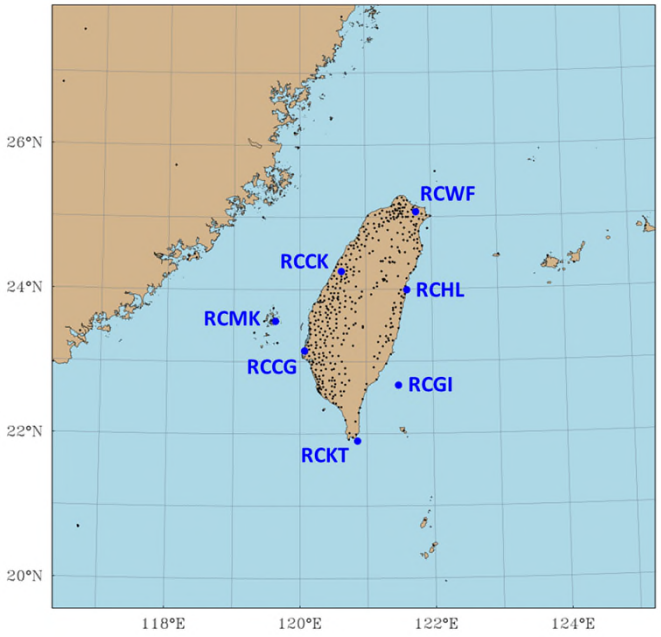


- ERA5 is not "true" hourly analysis data.
- This may <u>partly</u> explain why the 1-h Pangu-Weather model performs **worse than** its 24-h or 6-h versions.

CWA's regional reanalysis project

- **Goals:** **Deliver a high-resolution regional reanalysis dataset for the Taiwan area to support a variety of applications**, e.g.,
  - Climate studies
  - MLWP model training
- **Concepts and approaches:**
  - A high-resolution convection-allowing model
  - Rapid-update data assimilation of unique convective-scale observations
    → To provide (true) high-temporal-resolution data
  - A regional domain covering the vicinity of Taiwan
  - [For the 1st version] A period spanning ~10 years
- ***Nice!*** **We have CWA "RWRF" system.**

CWA RWRF (Radar WRF)



- Based on WRF and WRFDA
- **2-km resolution** for the Taiwan area
- **Hybrid 3DEnVar** coupled with RWRF-LETKF members
  - 30-minute cycle for radar DA
  - Hourly cycle for surface DA
- **6-hourly blending with the GFS downscaling**
- 13-hour forecasts initialized every hour
- **Radar data assimilation:**
  - Reflectivity (through hydrometeor retrievals) and radial velocity
- **Surface data assimilation:**
  - Temperature, humidity, winds

Development of RWRF-based regional reanalysis: Foundation and issues

- **Most things are based on a decade of development of CWA RWRF :**
  - Radar data assimilation *(蔡等人 2019, 大氣科學)*
  - Surface data assimilation *(Chen et al. 2020, WAF)*
  - Improvement of surface observation operators *(沈等人 2023, 天氣分析與預報研討會)*
  - Blending *(Jiang et al. 2021, TAO)*

- **For reanalysis purpose, two critical issues** in RWRF have been identified: **[both have been addressed!]**
  - The 6-hourly blending scheme* in RWRF can lead to **"discontinuous" hourly analysis fields**.

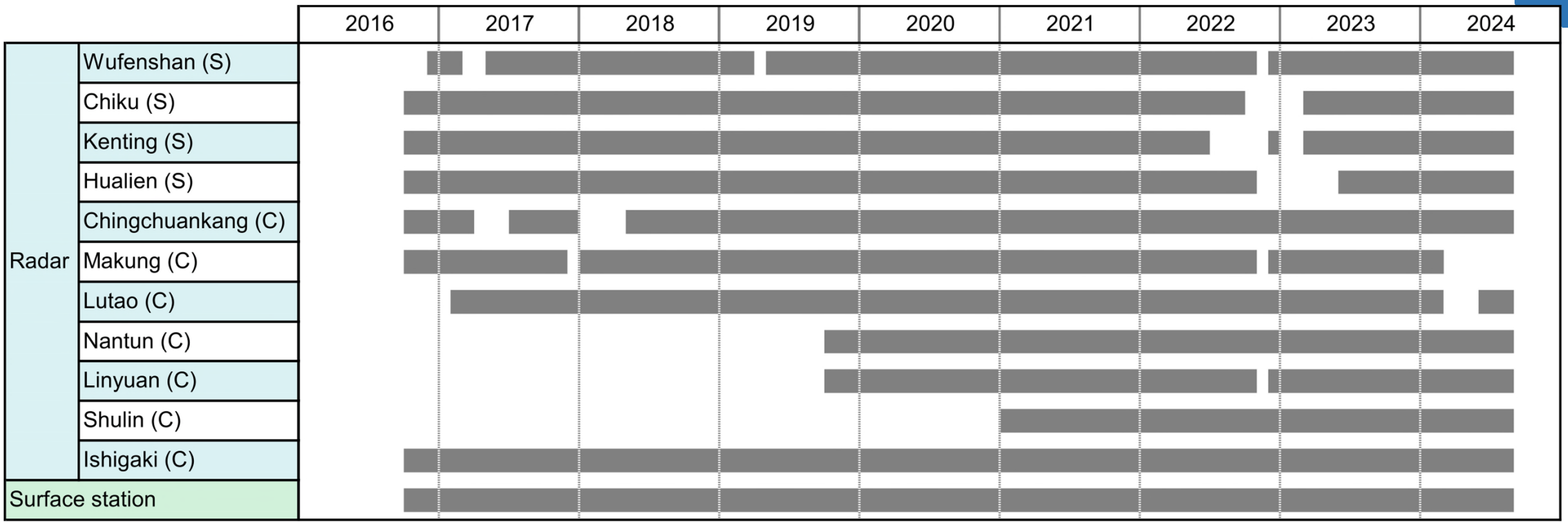
    ** To blend large-scale information (spatial scale > ~1200 km) from global analysis into the regional model fields.*

    → **Develop an "incremental blending update" method to improve the temporal smoothness.**
    [A2-2 沈彥志]
  - The current RWRF workflow is not quite suitable for long-term production.
    (e.g., disk usage, automation, and efficiency issues ...)
    → **Refactor the RWRF workflow scripts.**

Available Taiwan regional model analysis data & Plan for CWA RWRF-based regional reanalysis

	WRF dynamical downscaling from ERA5 Taiwan ReAnalysis Downscaling (**TReAD**) by TCCIP project	**RWRF operational analysis**	**RWRF-based reanalysis v1**	**RWRF-based reanalysis v2**
Model resolution	2 km			
Data assimilation method	(N/A)	Hybrid 3DEnVar	**3DVar**	**Hybrid 3DEnVar**
Assimilated observations		Surface observations, Radar radial winds, Radar reflectivity	Surface observations, Radar radial winds, Radar reflectivity	Surface observations, Radar radial winds, Radar reflectivity, GPS ZTD
Assimilation cycle length		1 hour (surface); 30 minutes (radar)	1 hour	1 hour or 30 minutes ?
Covered period	1979 – 2023	2018 – present	**2017 – 2026**	2017 – 2028 ?
Expected availability (year)	Available **NOW**	Available **NOW**	**2026**	**2029 ?**

Observation availability (first look)



Preliminary evaluations

Test period: **1 June – 7 June 2023**

	OP (operational RWRF at the time (2023))	**RA** (**reanalysis**: candidate version)	**NoDA** (similar to downscaling data like TReAD)	**ERA5**
WRF model version	OP441 (v3.8.1)	**OP51 (v4.4.2)**	**OP51 (v4.4.2)**	(ERA5)
Data assimilation method	Radar: 3DVar, every-30-min Surface: Hybrid, hourly	**3DVar, hourly**	—	
Surface observation operator	Default (via modified observation values)	**Improved** (via diagnosed surface variables)	—	
Digital filter initialization for surface DA	**V**	—	—	
Blending method	Traditional (Instantaneous update)	**Incremental blending update**	**Incremental blending update**	

Surface observation verification

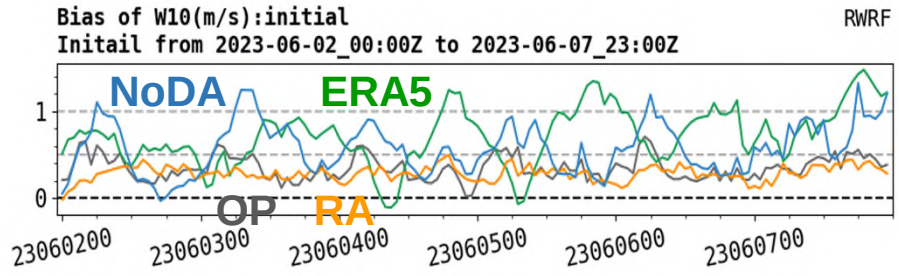
2 June – 7 June 2023,
analysis and 1–6 h forecasts

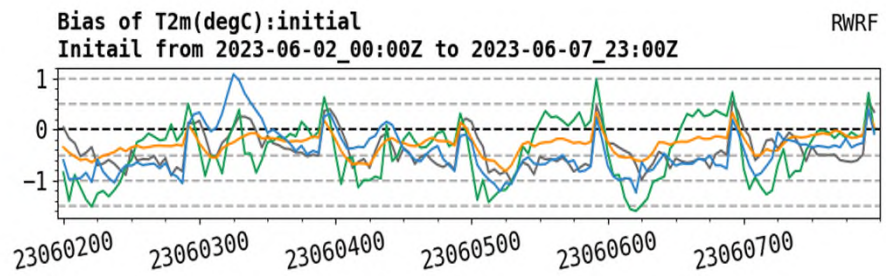
10-m wind speed (m/s)

2-m temperature (K)

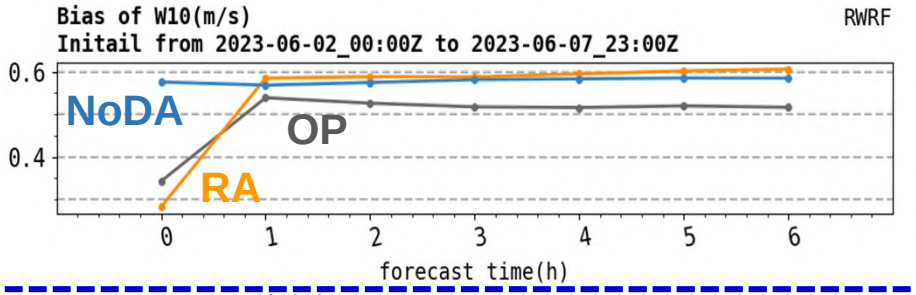
Bias

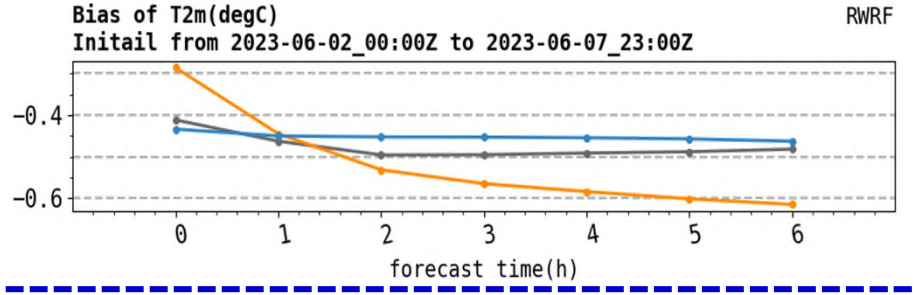
Time series of analysis





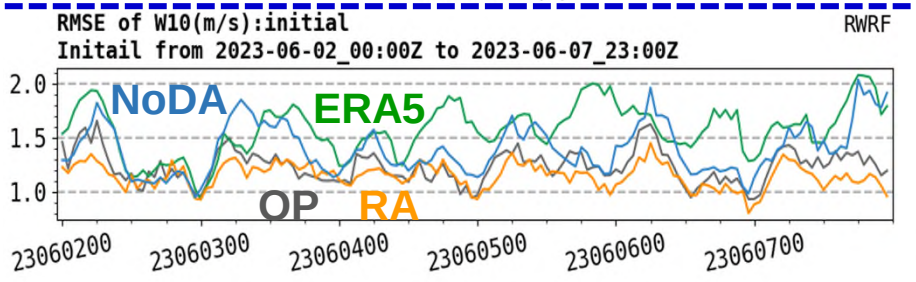
Average 0-6 h fcst

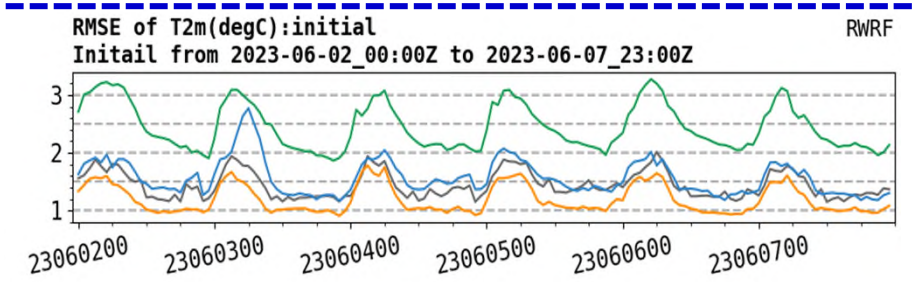




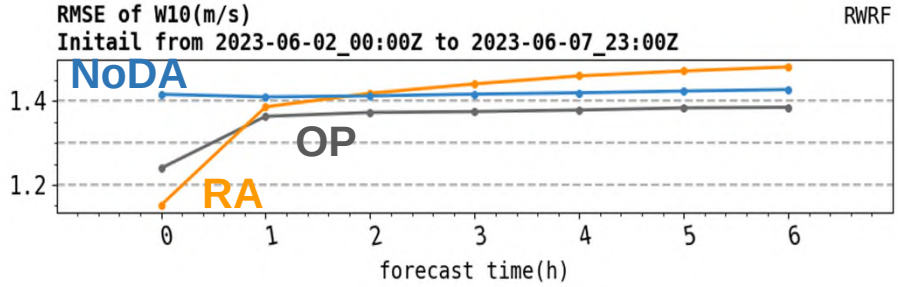
RMSE

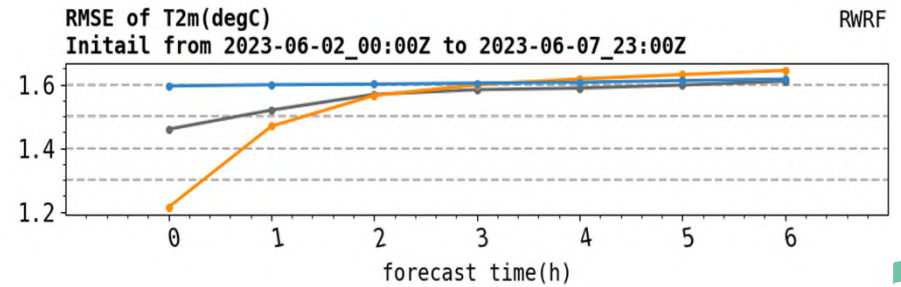
Time series of analysis





Average 0-6 h fcst



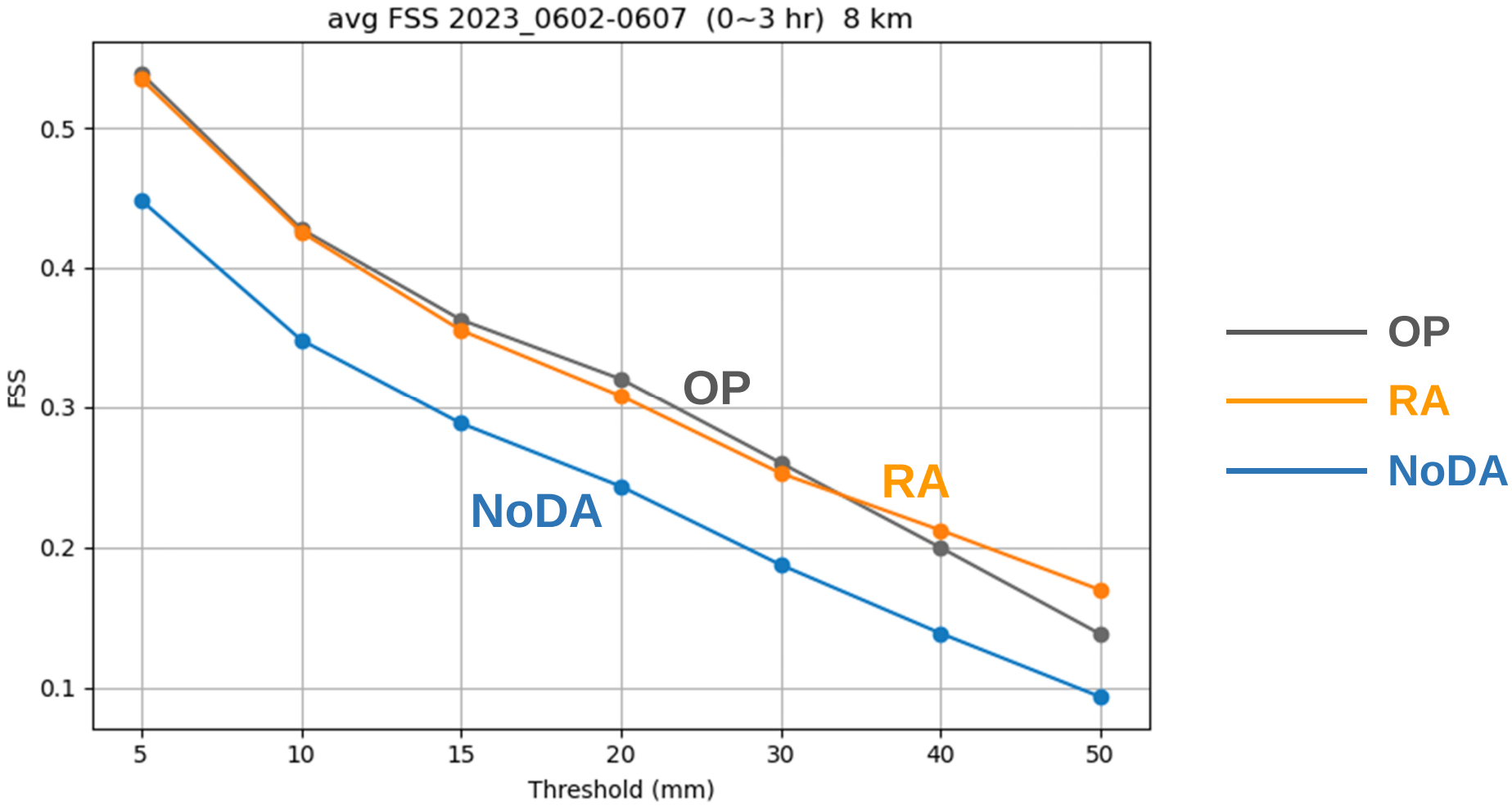


OP
RA
NoDA
ERA5

QPF verification

Fractions Skill Score (FSS)

2 June – 7 June 2023, 0–3 h accumulated rainfall, 8-km radius



Summary

- **CWA has started developing a regional reanalysis based on the RWRF system.**
  - 2-km resolution; hourly radar and surface data assimilation
  - To support a wide range of applications
  - Special emphasis on ensuring temporal smoothness **[A2-2 沈彥志]**
- **Timeline:**
  - Configurations yet to be finalized before starting production
  - Targeted for “version 1” release by the end of the next year (2026)

  ** For AI/ML work TODAY, there have been choices of available regional training data (e.g., RWRF operational analysis, TReAD)*
- **Preliminary evaluation in test periods:**
  - Better than ERA5 and “NoDA” (downscaling-like run)
  - Roughly comparable to the operational analysis, with better fit to some observations and improved temporal smoothness.

Backup slide

Compilation of RWRF's operational data

- **2-km resolution, hourly analysis**
- **Period : 2018 – 2025 (7+ years)**

Variable list:

Vertical levels (31):
- 20, 30, 50, 70, 100, 125, 150, 175, 200, 225, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000 hPa

3D variables (11):
- umet, vmet, tk, z, wa, QVAPOR, QRAIN, QSNOW, QGRAUP, QCLOUD, QICE

2D variables (14):
- PSFC, slp, umet10, vmet10, T2, Q2, rh2, td2, SST, pw, PBLH, RAINNC, SWDOWN, OLR

Fixed variables (4):
- XLAT, XLONG, HGT, LANDMASK

Analysis / forecast times (8):
- Analysis (wrfinput), 0, 1, 2, 3, 4, 5, 6-hour forecasts (wrfout)

- Vertical interpolation, computation of physical variables, and data format conversion are performed in the data compilation process.
- About 5% of the data are missing due to operational or archival failures.