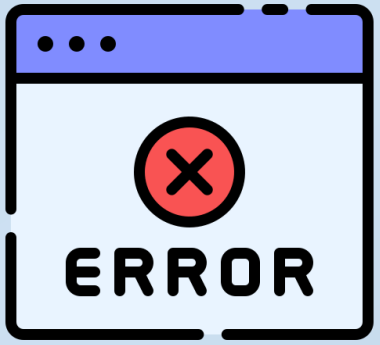
6th WP-MIP Project Meeting

Coordinators' Update
15 June 2026

Agenda

- WP-MIP Update (Coordinators)
  - Problems with forecast contribution uploads to ECCC
  - JAG-AI feedback
  - Reading “ragged” WP-MIP data
  - SP1 TC tracking
- Verification theme updates (theme leads; focus on falsification?)
- Verification Quick-Looks:
  - Extreme Weather Bench (Amy)
  - Spectrally decomposed noise and information energies (Britta and Sabrina)
  - HiRA temperature scorecard (Marion)

Forecast Upload Problems



- At the end of May, contributors started to encounter "input/output" errors when they uploaded forecasts
- Data files appear to be the right size, but appear to be corrupted when read
- This problem is not unique to WP-MIP: other groups using the ECCC servers have reported similar issues
- IT support from Shared Services Canada solved the problem on12 June WP-MIP tracking at https://github.com/WP-MIP/.github/issues/51

Downloads of data from the WP-MIP archive are **not** affected

JAG-AI Feedback

The WIPPS Secretariat asked WP-MIP to present progress to JAG-AI at the end of March.

The group was very pleased with WP-MIP progress and provided a few recommendations:

- Collect information from contributors about computing requirements
- Account for sector-specific needs in verification activities
- Plan future extension to ensembles and probabilistic verification

WMO Joint Advisory Group on AI

Advice, oversight, coordination and monitoring of WMO AI activities.

Reading "Ragged" Datasets

Combination of core and subproject data leads to some "ragged" contributions, where different forecast ranges are available for different dates.

Example: ECMWF forecasts extend to 240 h on core dates and 336 h on SP1 dates

Default conversion from an Earthkit FieldList to xarray fails with a shape error.

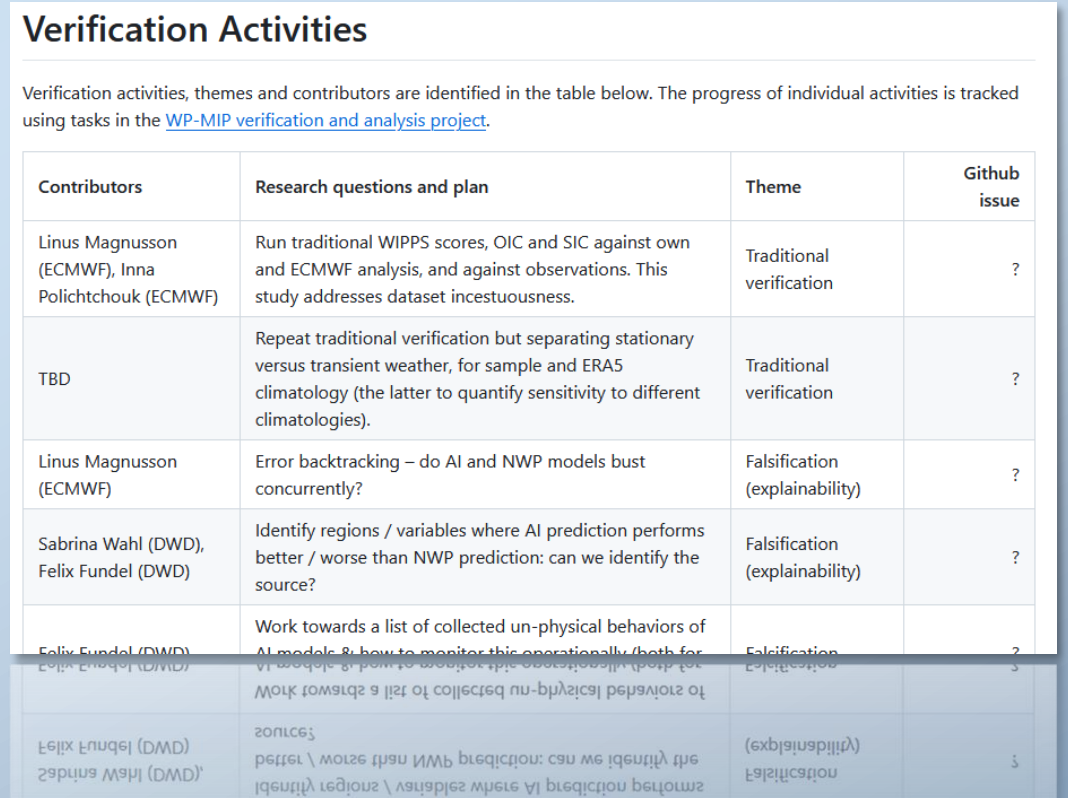
The original score.py example in the scores package has been updated to handle these datasets using the `allow_holes` argument:

```
0,7 @@ def compute(self, scores):
 fname = '_'.join([self['field'], self['level'], self['stream'], self['centre'],
                   self['ftype'], self['version'], 'pl', month])+'.grib2'
 fpath = os.path.join('forecasts', self['stream'], fname)
 ds = ek.data.from_source('file', fpath).to_xarray()
 ds = ek.data.from_source('file', fpath).to_xarray(allow_holes=True)
 for init, this_init in ds.groupby("forecast_reference_time"):
     init_dt = datetime.fromisoformat(str(this_init.forecast_reference_time.values[0]))
     if init_dt not in self['icdates']:
```

Verification Theme Updates

- Traditional verification
- Falsification:
  - Explainability
  - Physical consistency
- Spatial methods
- Evaluation of extremes
- Tropical meteorology
  - Tropical cyclones (incl. SP1)
  - Everything else (incl. SP2)
- Case and regional studies

Remember to add your information to the verification table to help theme leads to track and coordinate activities.



Verification Activities

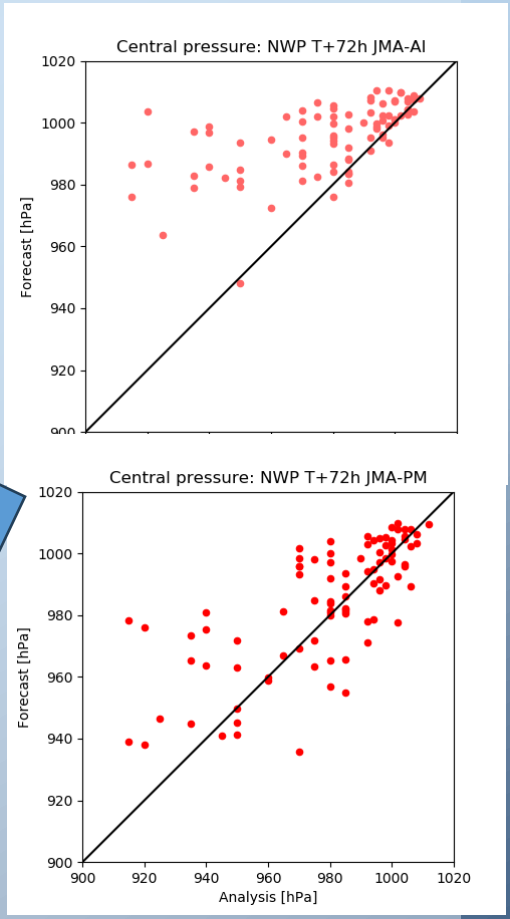
Verification activities, themes and contributors are identified in the table below. The progress of individual activities is tracked using tasks in the WP-MIP verification and analysis project.

Contributors	Research questions and plan	Theme	Github issue
Linus Magnusson (ECMWF), Inna Polichtchouk (ECMWF)	Run traditional WIPPS scores, OIC and SIC against own and ECMWF analysis, and against observations. This study addresses dataset incestuousness.	Traditional verification	?
TBD	Repeat traditional verification but separating stationary versus transient weather, for sample and ERA5 climatology (the latter to quantify sensitivity to different climatologies).	Traditional verification	?
Linus Magnusson (ECMWF)	Error backtracking – do AI and NWP models bust concurrently?	Falsification (explainability)	?
Sabrina Wahl (DWD), Felix Fundel (DWD)	Identify regions / variables where AI prediction performs better / worse than NWP prediction: can we identify the source?	Falsification (explainability)	?
Felix Fundel (DWD)	Work towards a list of collected un-physical behaviors of AI models & how to monitor this operationally (both for	Falsification	?

Current progress from SP1: Tropical Cyclones

- The **WGNE-based TC verification** (track and central pressure) are progressing using SP1 and CORE datasets as of June 2026
  - TC track using mslp (mean sea level pressure)
  - Contributions from 9 centres (CMA, CMC, DWD, ECMWF, GFDL, JMA, KIAPS, NCEP, UKMO) and 17 models including AI, HY, and PM
    - Tracking files are being available soon.
    - A sample csv file (SP1 over Western North Pacific area) and a python script are available
    - Based on DataFrame structure, easily readable using Excel, Python (pandas required), etc…

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	region	tcnum	tcname	inittime	validtime	leadtime	fcst_lat	fcst_lon	best_lat	best_lon	fcst_pc	best_pc	center
2	NWP	2401	EWINIAR	2024052400	2024052400	0	8.7	129.1	8.5	128.3	1006.7	1006	JMA-PM
3	NWP	2401	EWINIAR	2024052400	2024052400	0	8.3	128.3	8.5	128.3	1006.6	1006	CMC-PM
4	NWP	2401	EWINIAR	2024052400	2024052400	0	8.8	127.9	8.5	128.3	1007	1006	ECMWF-PM
5	NWP	2401	EWINIAR	2024052400	2024052400	0	8	128	8.5	128.3	1006.6	1006	GFDL-PM
6	NWP	2401	EWINIAR	2024052400	2024052400	0	7.5	128.8	8.5	128.3	1004.8	1006	NCEP-PM
7	NWP	2401	EWINIAR	2024052400	2024052400	0	8.7	129.1	8.5	128.3	1006.7	1006	JMA-AI
8	NWP	2401	EWINIAR	2024052400	2024052400	0	8.3	128.3	8.5	128.3	1006.8	1006	CMC-AI
9	NWP	2401	EWINIAR	2024052400	2024052400	0	8.3	128.3	8.5	128.3	1006.6	1006	CMC-HY
10	NWP	2401	EWINIAR	2024052400	2024052400	0	8.7	128	8.5	128.3	1007.1	1006	ECMWF-AI
11	NWP	2401	EWINIAR	2024052500	2024052500	0	12.5	124.5	12.3	123.8	1005.6	1006	JMA-PM



- **The GFDL vortex tracker outputs *version 1.0*** is planned to release in late June.
  - Issues for ECCC and UKMO models in *version 0.0* have been detected and the related files were deleted.
  - Updates for ECCC, UKMO, NCEP models, and new data for DWD, KMA, RAS, and INPE will be included.

Verification Quicklooks

Extreme Weather Bench (Amy)
Noise and Information (Britta & Sabrina)