香港天文台
HONG KONG OBSERVATORY

Object-based verification of NWP model performance regarding the wind structure of Super Typhoon Hato

Wai-Kin WONG[1], WC Woo[1], KC Au-Yeung [1] and YC Wong[2]

[1]Hong Kong Observatory

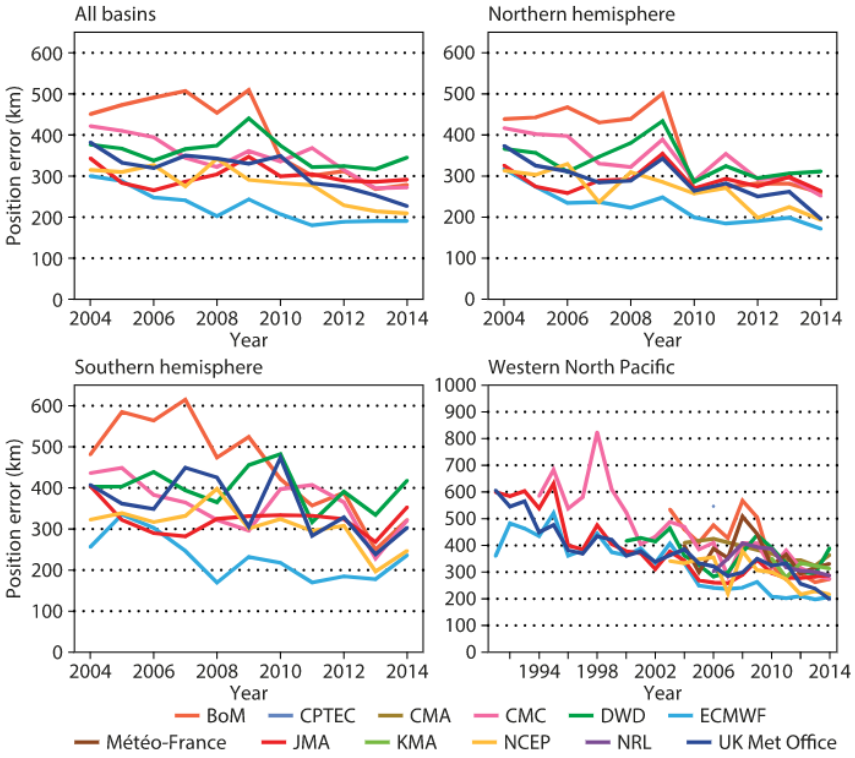
[2]The Chinese University of Hong Kong

ESCAP/WMO Typhoon Committee Technical Conference (TECO)
26 – 27 February 2018
Hanoi, Viet Nam

香港天文台
HONG KONG OBSERVATORY

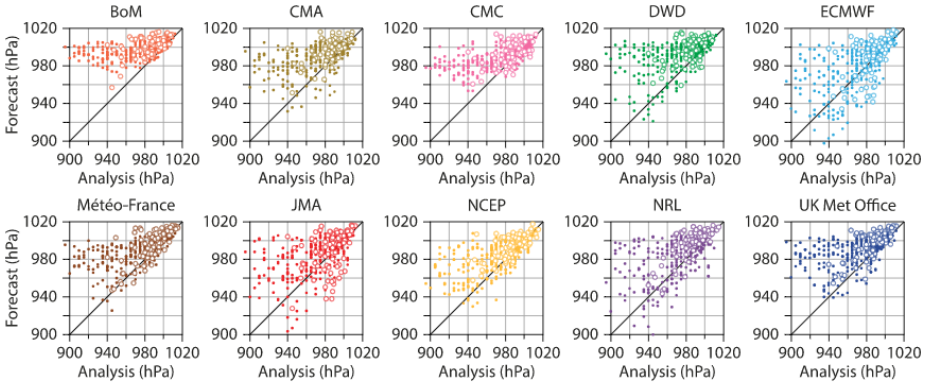
Common TC Verification Metrics

Position



Timeseries of TC position forecast errors. The charts show the evolution of annual average TC position errors from 3-day forecasts for the models participating in the intercomparison project for all TC basins (top left), the northern hemisphere (top right), the southern hemisphere (bottom left) and the western North Pacific (bottom right).

Intensity

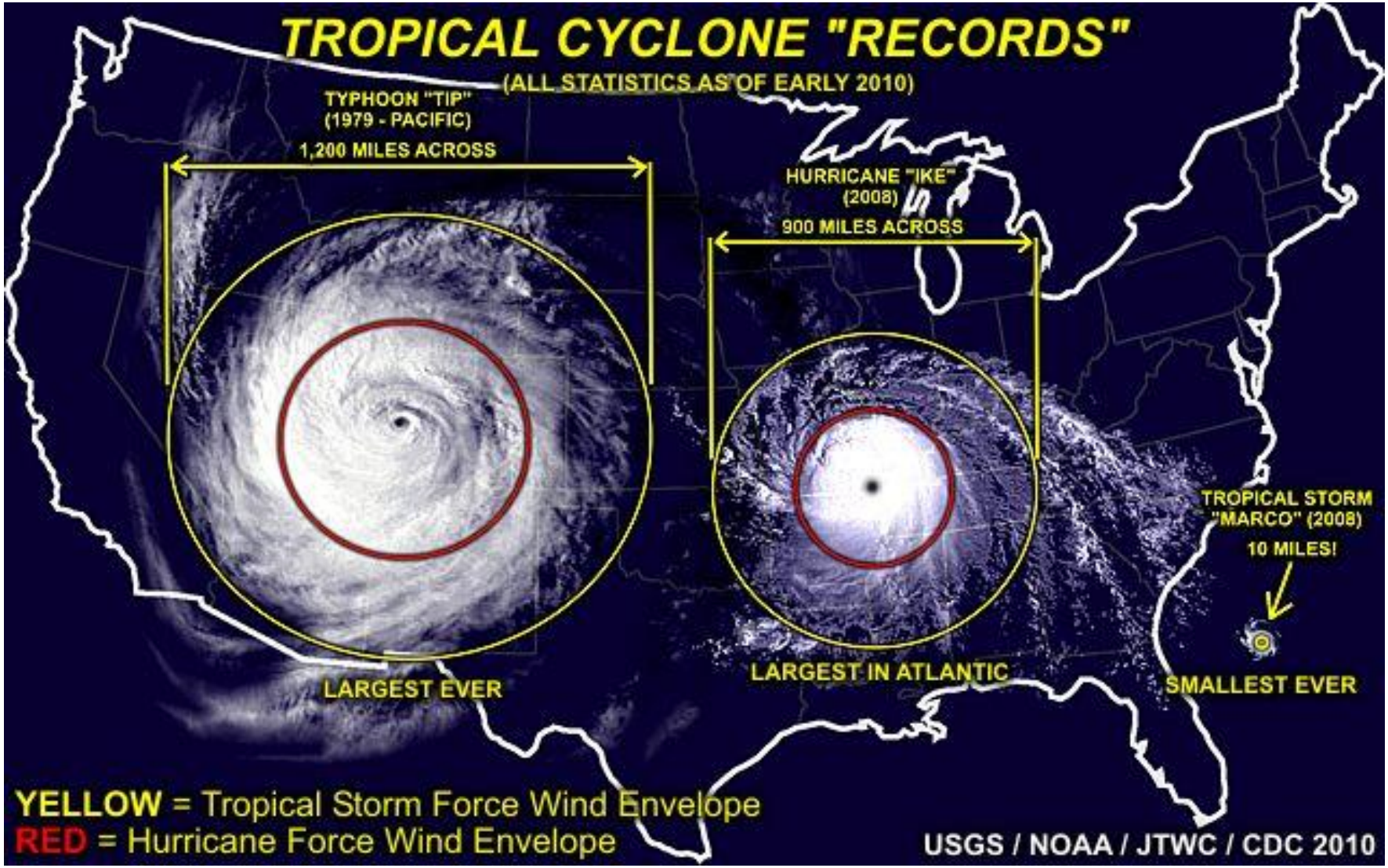


Intensity biases in TC forecasts. Scatter plots of analysed (best track) versus 3-day forecast minimum sea level pressure, indicating how the different models share similar TC intensity biases especially for TCs analysed to have very low minimum pressure. The verification period is 2012–2014.

Source: Working Group on Numerical Experimentation (WGNE), jointly established by the World Climate Research Programme (WCRP) Joint Scientific Committee (JSC) and the World Meteorological Organization's Commission for Atmospheric Sciences (CAS)

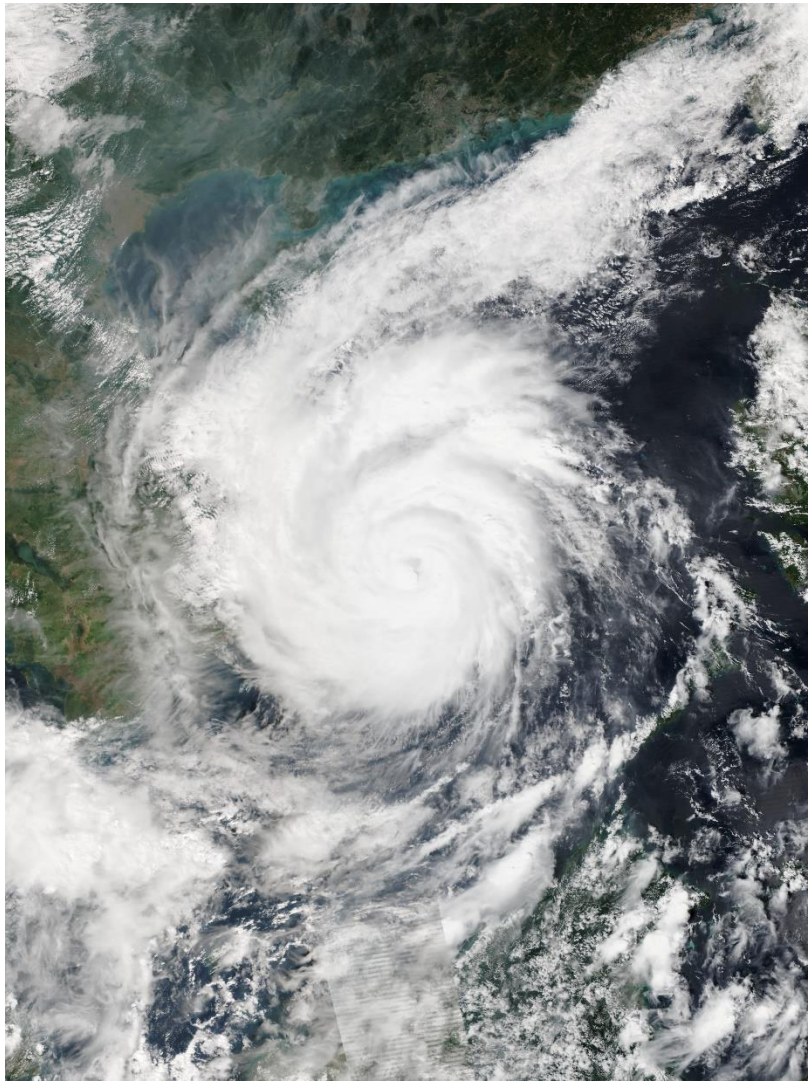
香港天文台
HONG KONG OBSERVATORY

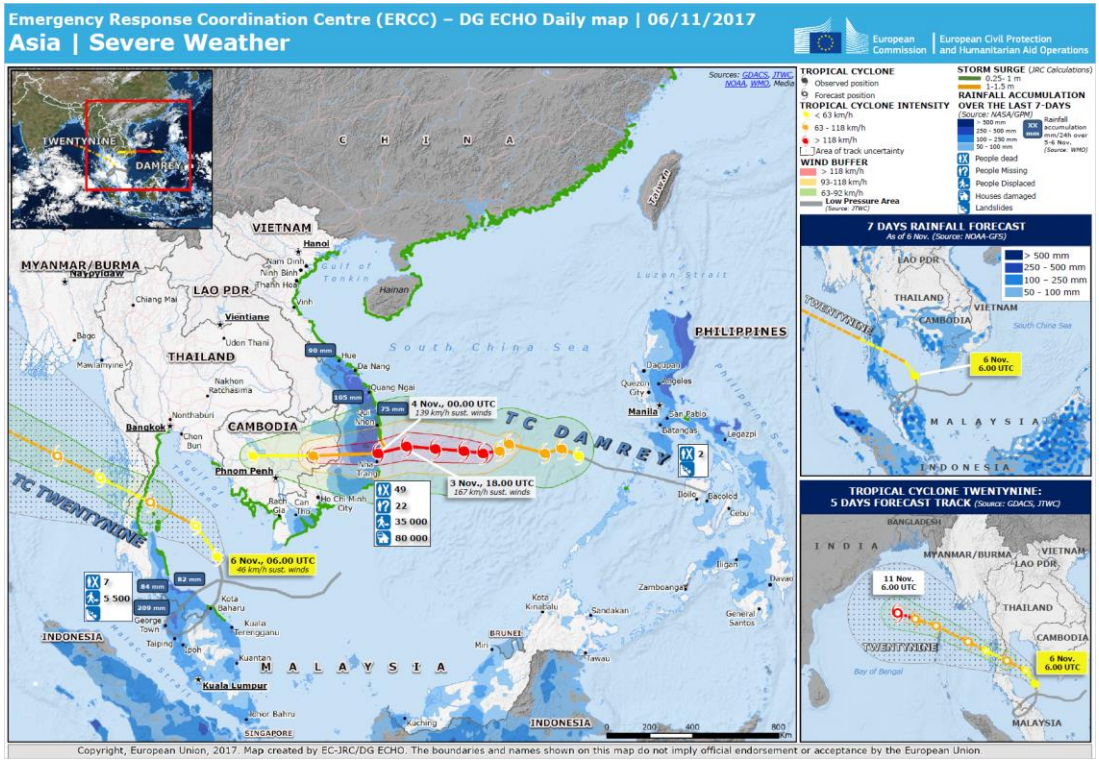
But TC comes in all sizes…



香港天文台
HONG KONG OBSERVATORY

Typhoon Damrey approaching Vietnam on 3 Nov 2017





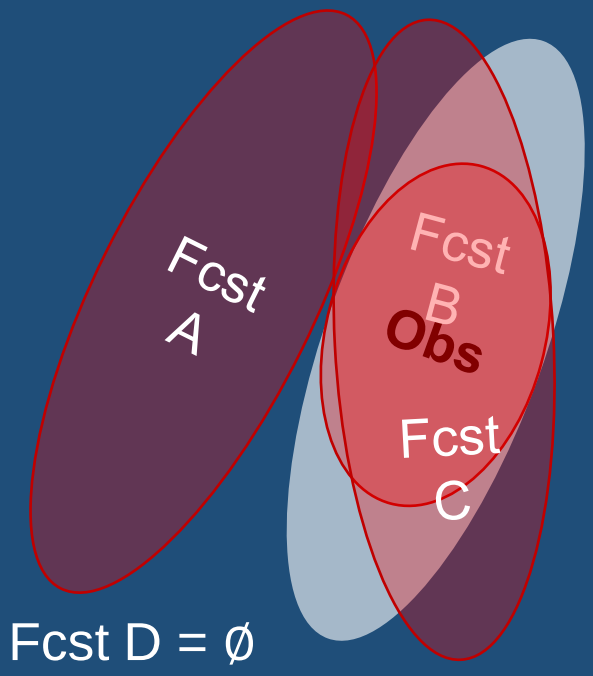
香港天文台
HONG KONG OBSERVATORY

Impact Based Warnings

- Multi-levels Warnings based on local wind strength, for example:
  - No. 1 / 3 / 8 / 9 / 10 in Hong Kong & Macao
  - Blue / Yellow / Orange / Red in Mainland China
  - No. 1 - 5 in the Philippines
  - Watch / Warning in Australia
  - Stage 1 - 4 in India
  - Emergency Warning / Warning / Advisory in Japan

- Local wind strength depends not only on location and intensity of TC, but also wind structures

Object-Based Verification



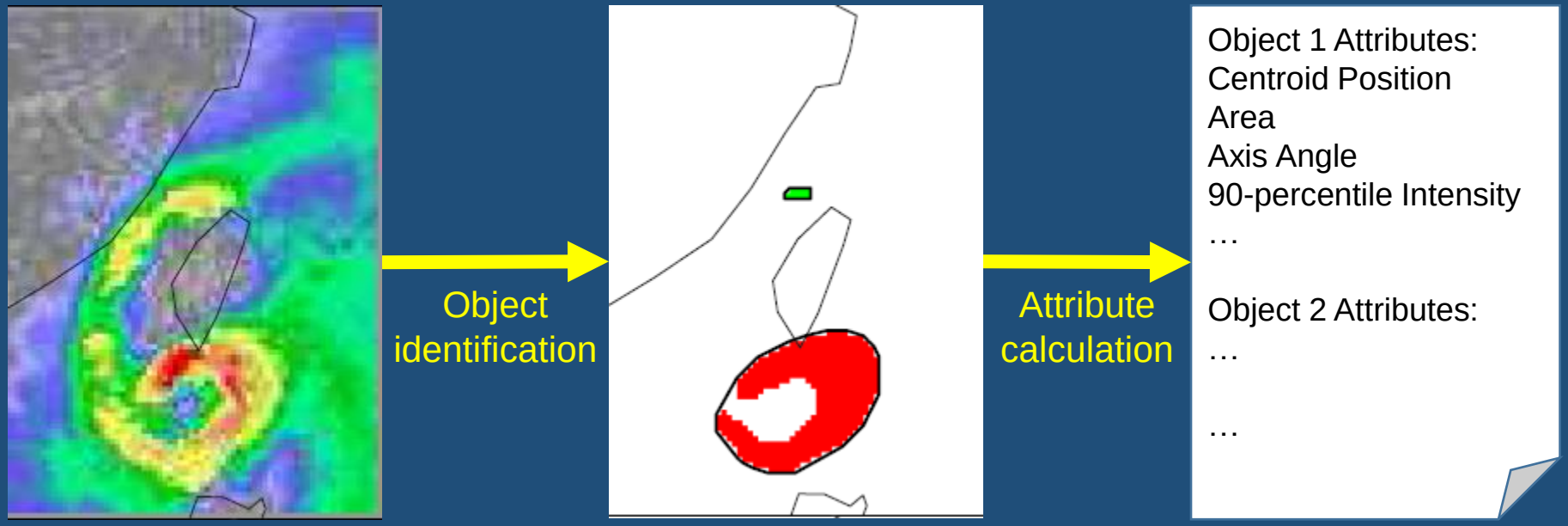
	Subjectively	CSI
Fcst A	Inaccurate Position	0
Fcst B	Inaccurate Size	~0.5
Fcst C	Inaccurate Axis Angle	~0.5
Fcst D	Miss	0

Object-based verification is to provide:

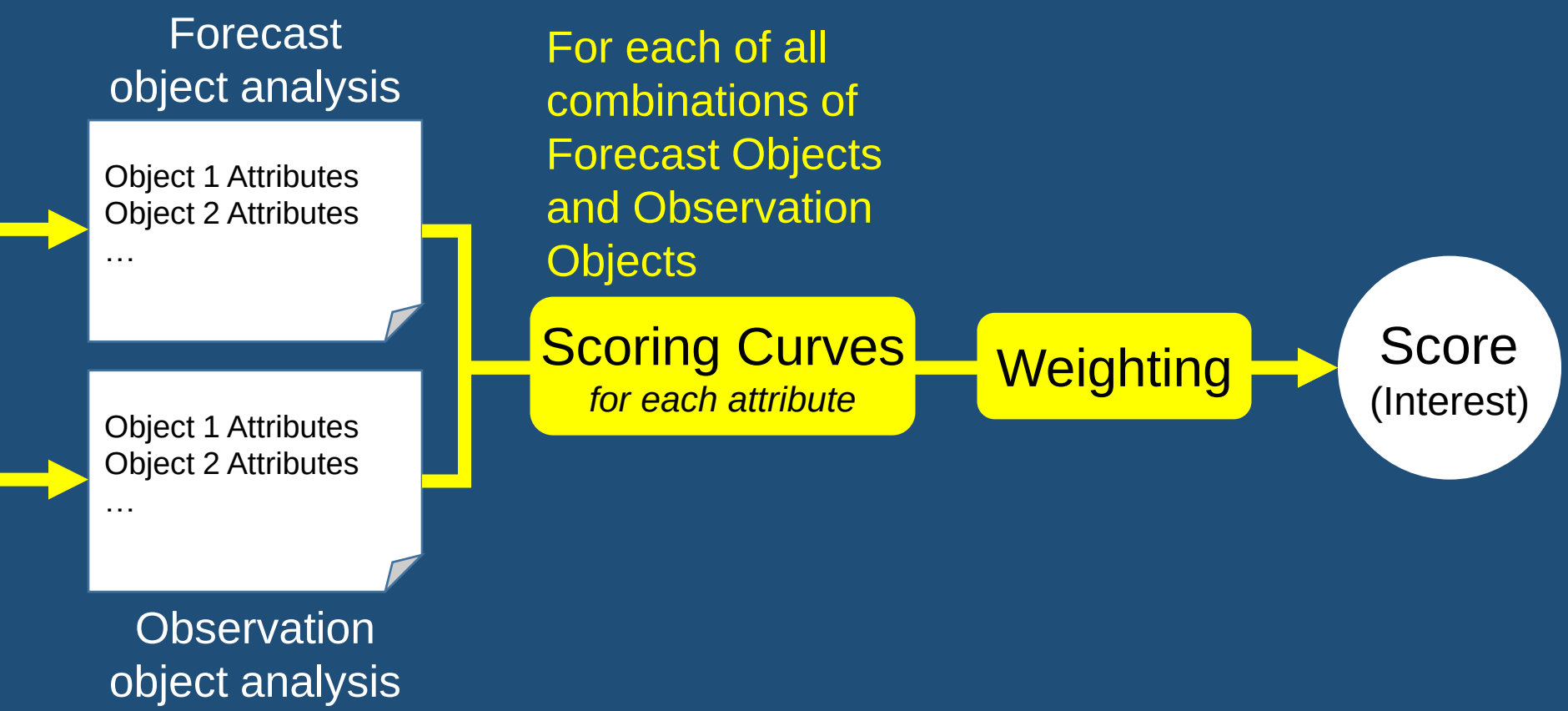
- ***Performance for each attribute*** (e.g. position, angle, intensity, etc.)
- ***An overall score*** considering all attributes with weighting

MODE *(Method for Object-Based Diagnostic Evaluation)*

- NCAR - Developmental Testbed Center (DTC)
  - MET *(Model Evaluation Tools)* verification package
    - MODE
    - https://dtcenter.org/met/users/

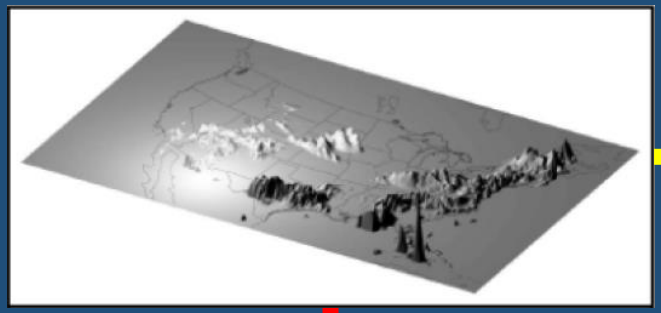


MODE *(Method for Object-Based Diagnostic Evaluation)*

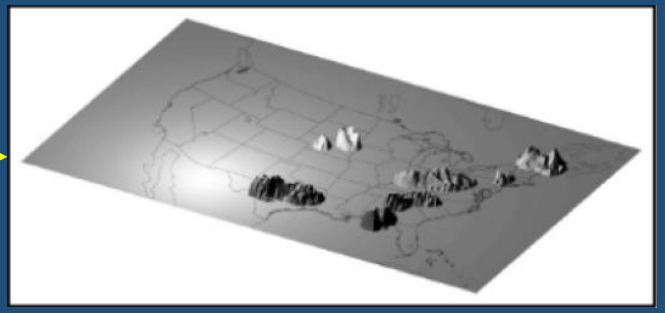


Object Identification

Raw Field



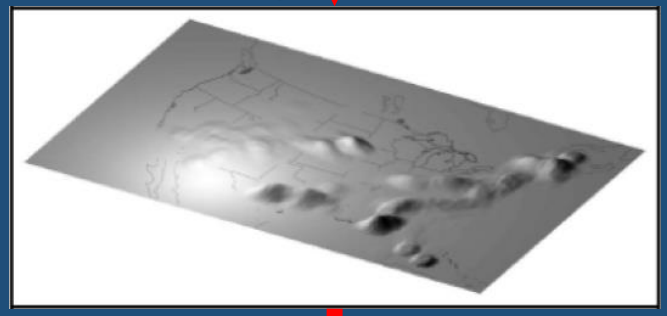
Masking



Objects

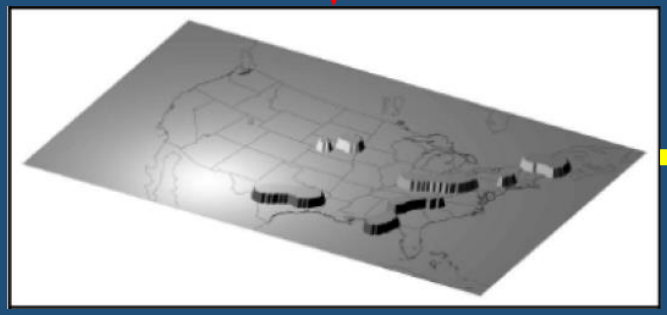
Convolution
(Parameter: Convolution Radius)

Convolved Field



Mask Creation
(Parameter: Threshold)

Mask



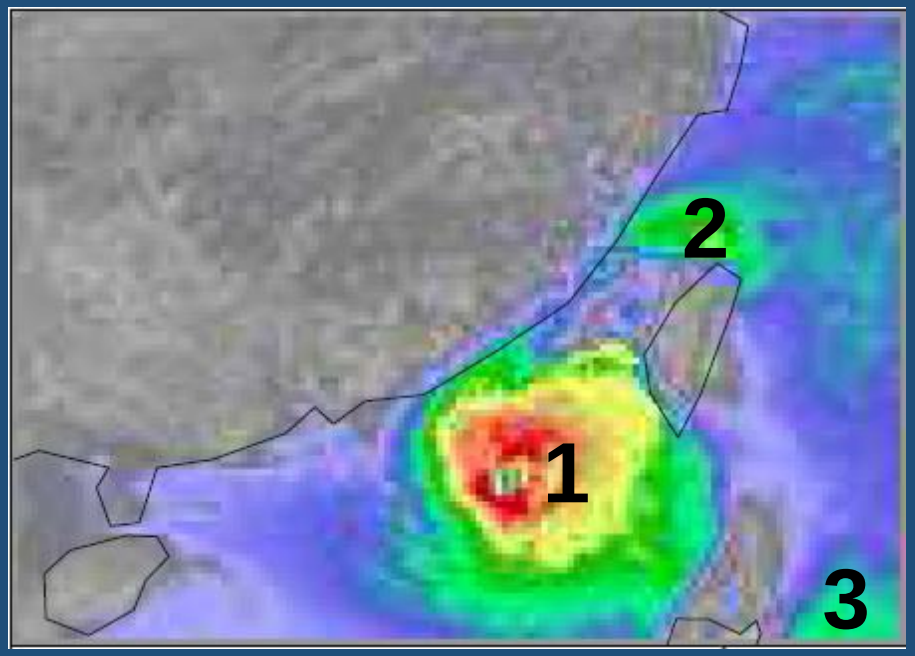
Why Convolution?

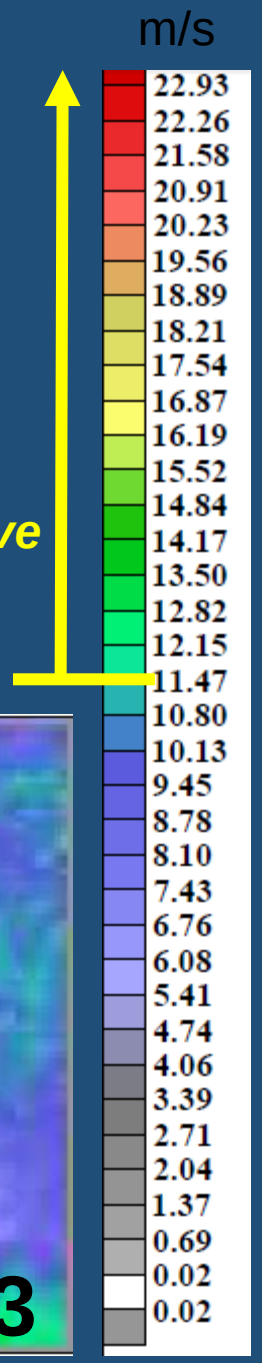
How many pieces of bread here?

3. Normally we don't count the crumbs.

Strong Wind or above

ECMWF 08-21 12Z t+24 Forecast





How many Strong Wind (or above) objects?

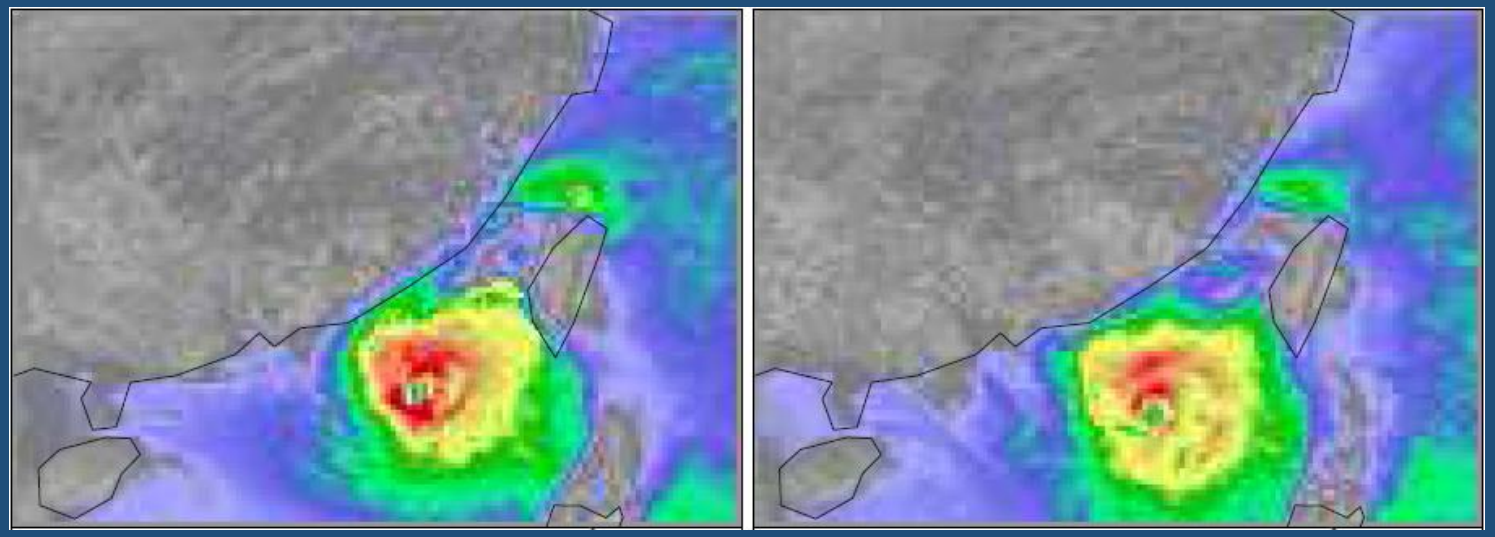
3

Why Convolution?

ECMWF 08-21 12Z t+24

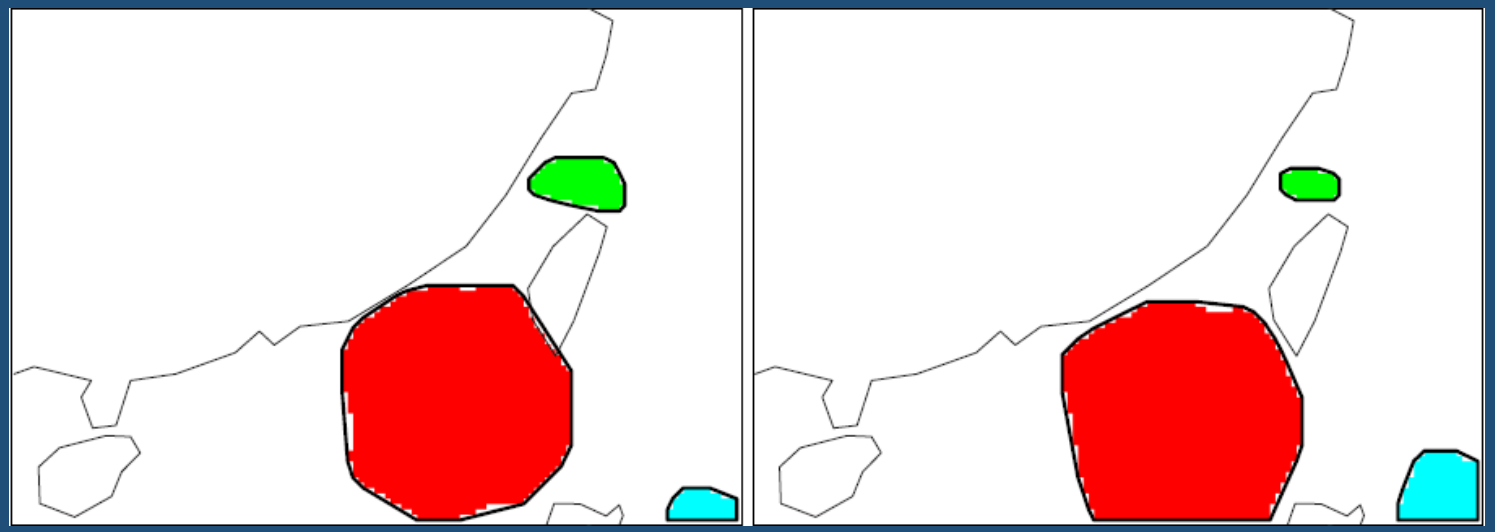
ECMWF 08-22 12Z Analysis

Raw Field



Strong Wind Objects

(convolution radius = 5)

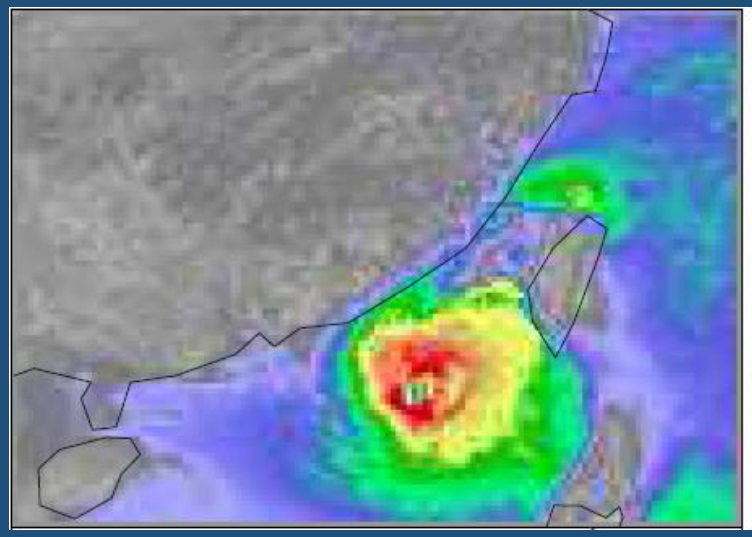


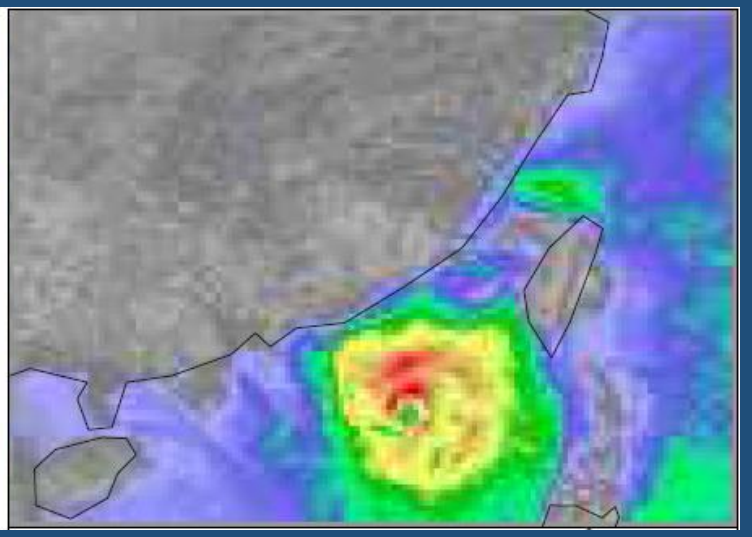
Why Convolution?

ECMWF 08-21 12Z t+24

ECMWF 08-22 12Z Analysis

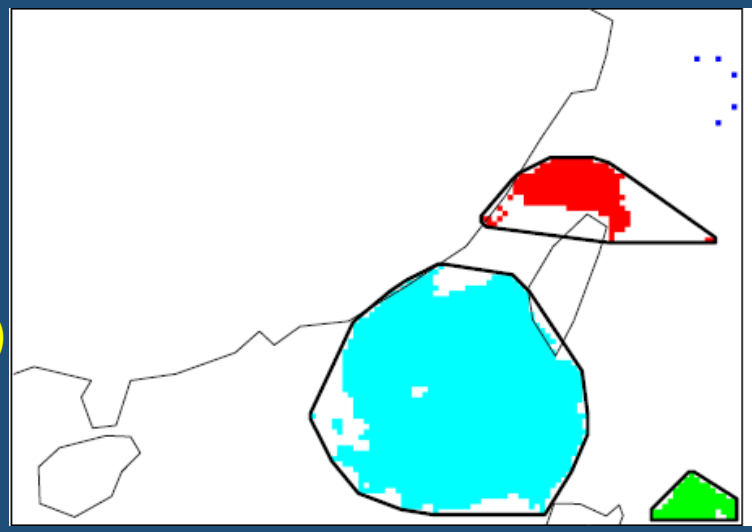
Raw Field

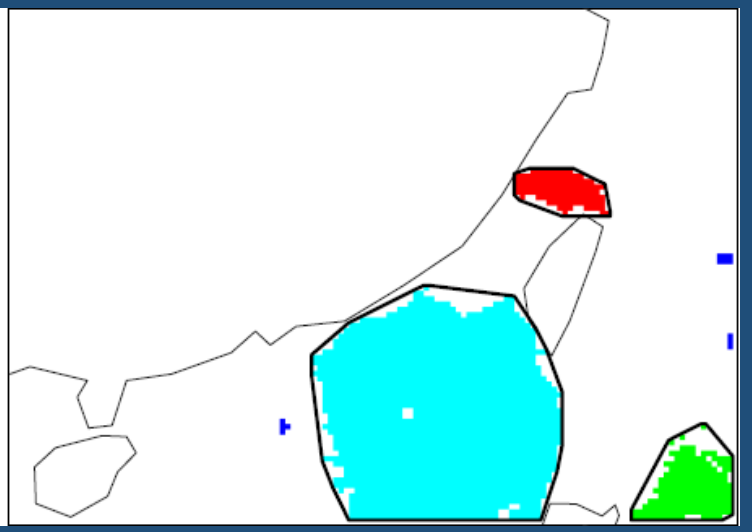




Strong Wind Objects

(no convolution)





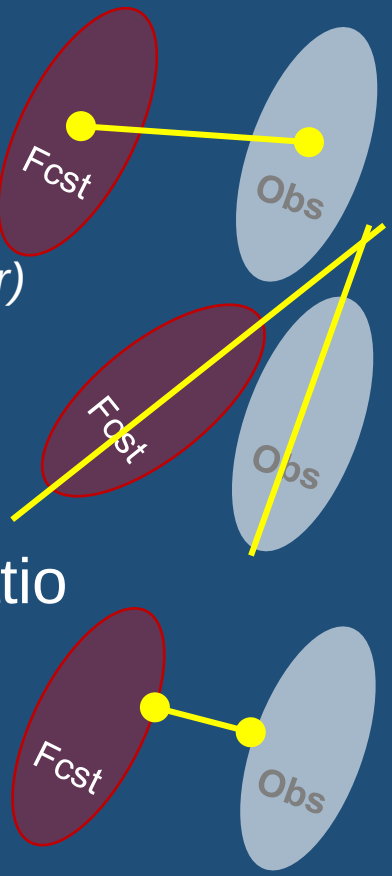
Attributes

Single Object Attributes

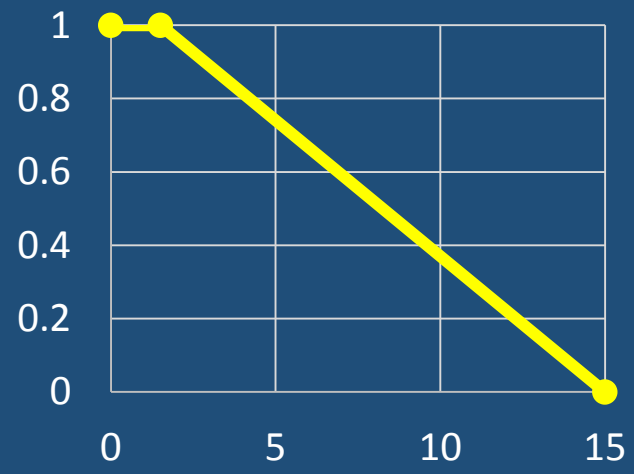
- Centroid Position →
- Area →
- Axis Angle →
- Intensity Percentile →
- *and more …*

Object Pair Attributes

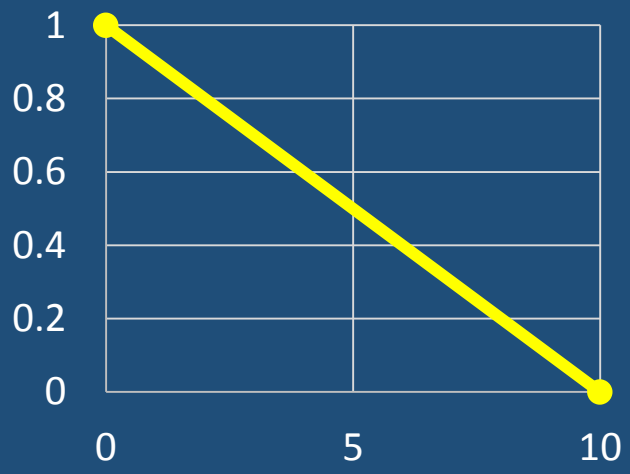
- Centroid Distance
- Area Ratio *(Smaller/Larger)*
- Axis Angle Difference
- Intensity Percentile Ratio *(Smaller/Larger)*
- Boundary Distance
- Interception Area Ratio *(Interception Area / Smaller of the {Fcst,Obs} Area)*



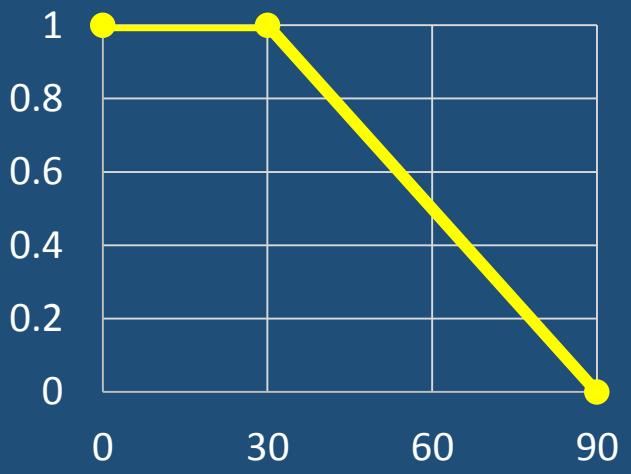
Scoring Curves & Weighting (Defaults in MODE)



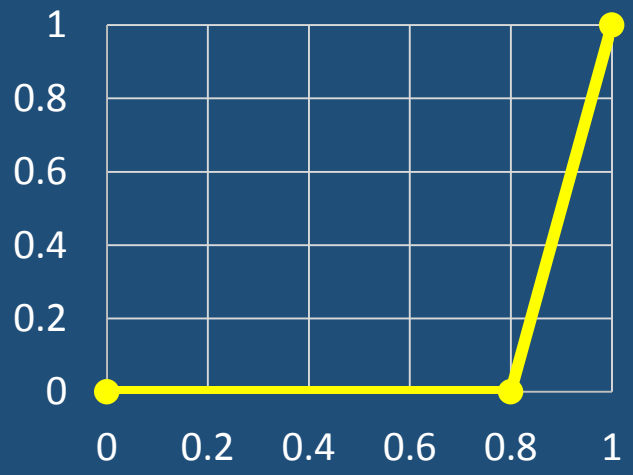
Centroid Distance (Grid)
Weighting: 2



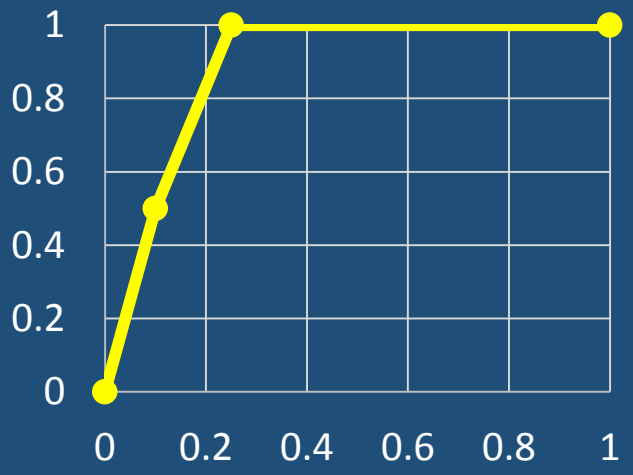
Boundary Distance (Grid)
Weighting: 4



Axis Angle Difference (degree)
Weighting: 1



Area Ratio
Weighting: 1



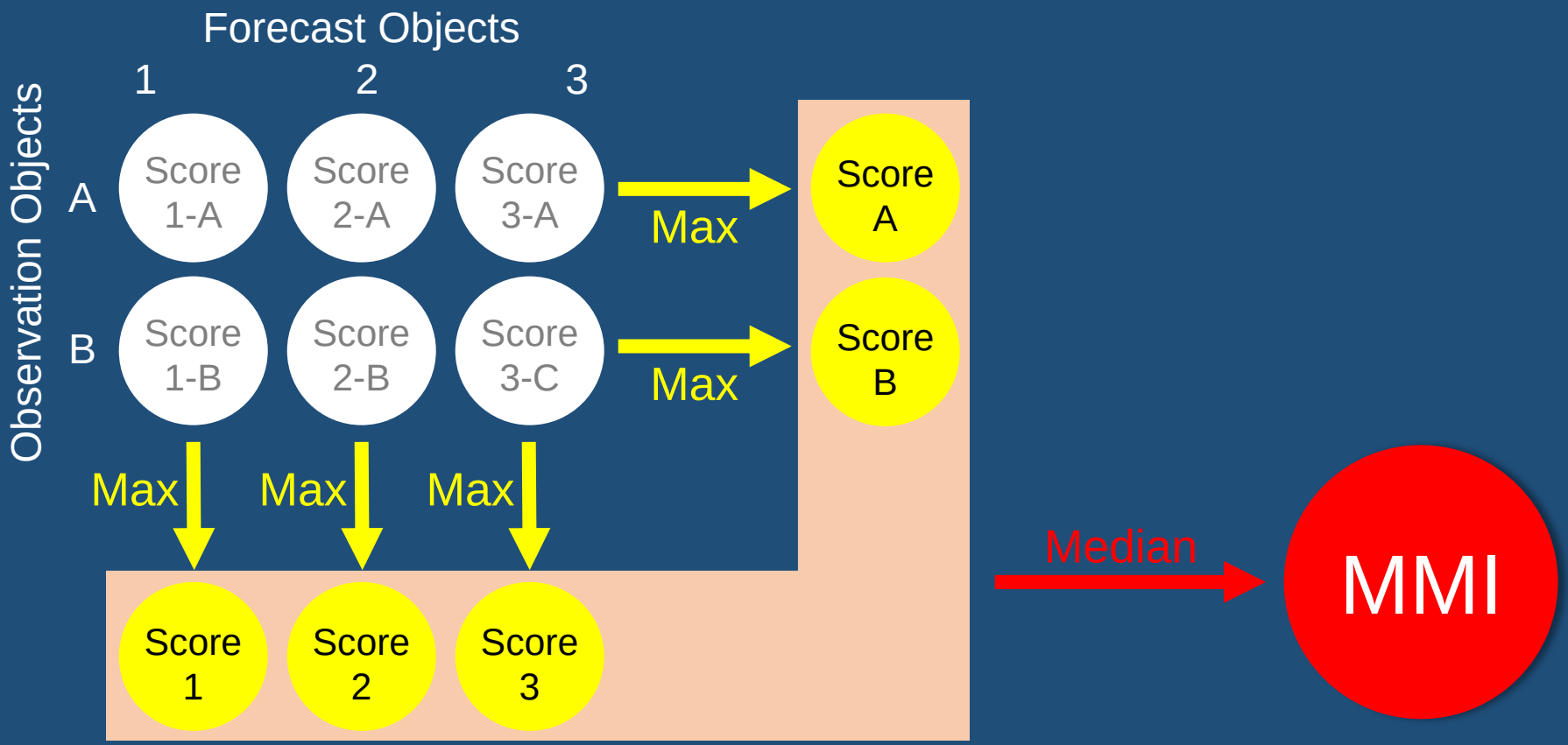
Interception Area Ratio
Weighting: 2

MMI (Median of Maximum Interest)

- A score for the whole forecast considering all objects

Example

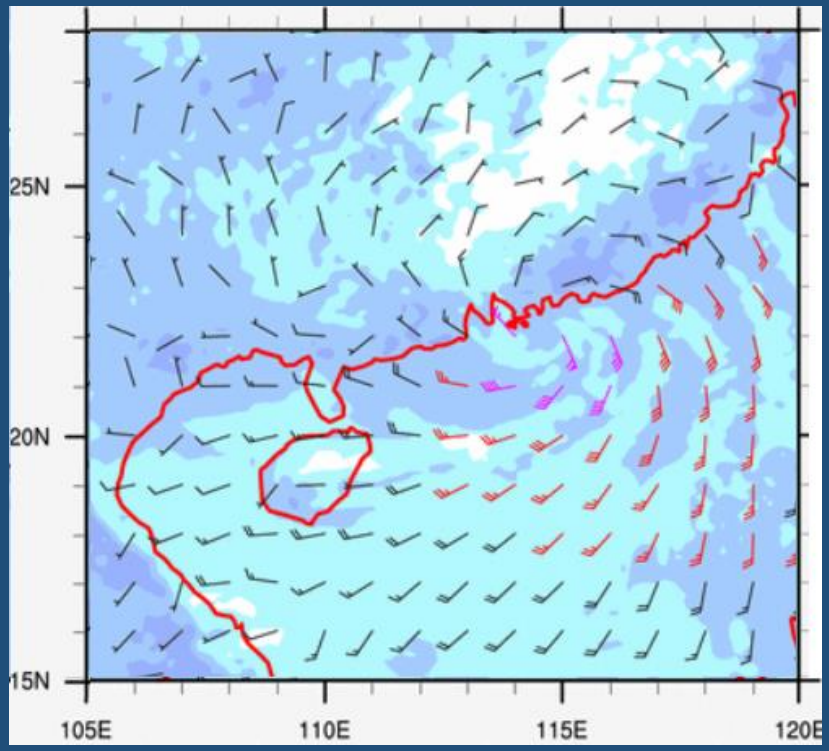
Forecast Objects: 1, 2, 3
Observation Objects: A, B



Data adopted in this study:

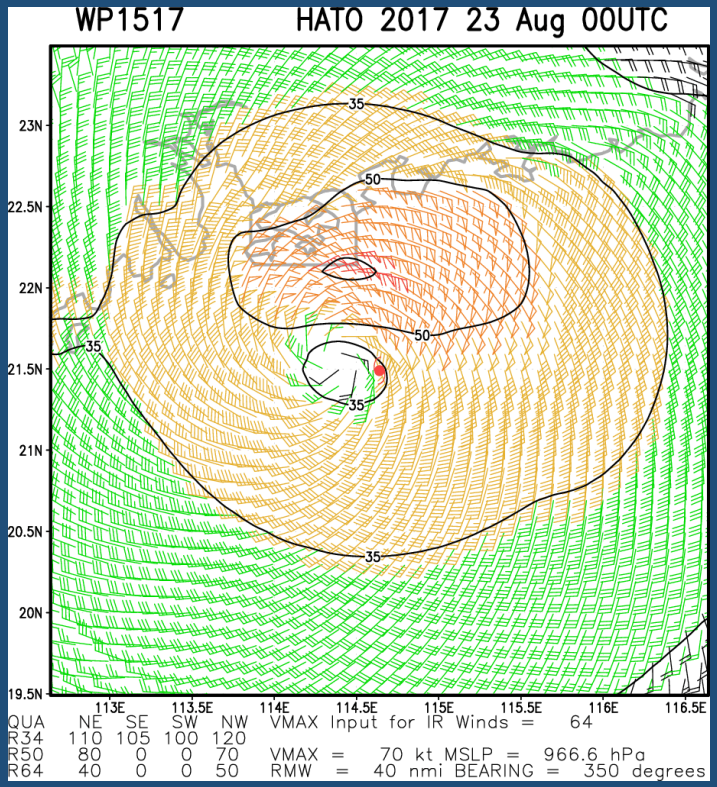
Forecast

Outputs from NWP models (ECMWF, JMA, NCEP, UKMO)



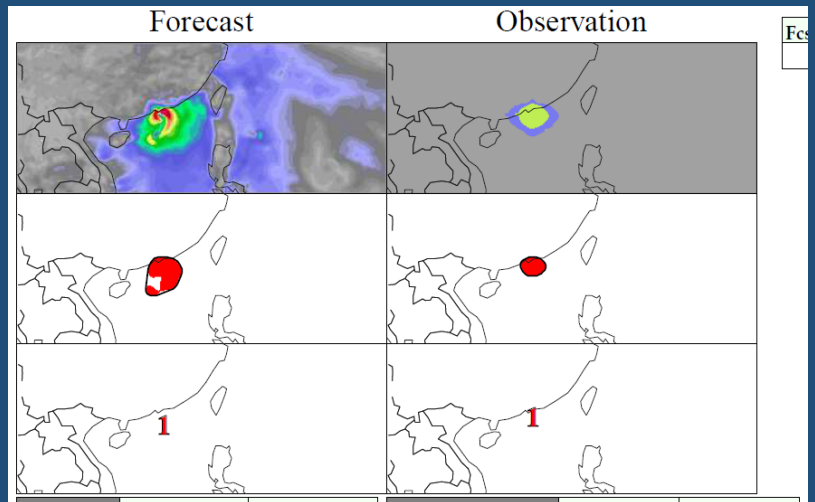
Observation

NOAA Multi-Platform Tropical Cyclone Wind Analysis



Base Time: 2017-08-21 12Z, 36-hour f/c
Valid Time: 2017-08-23 00Z

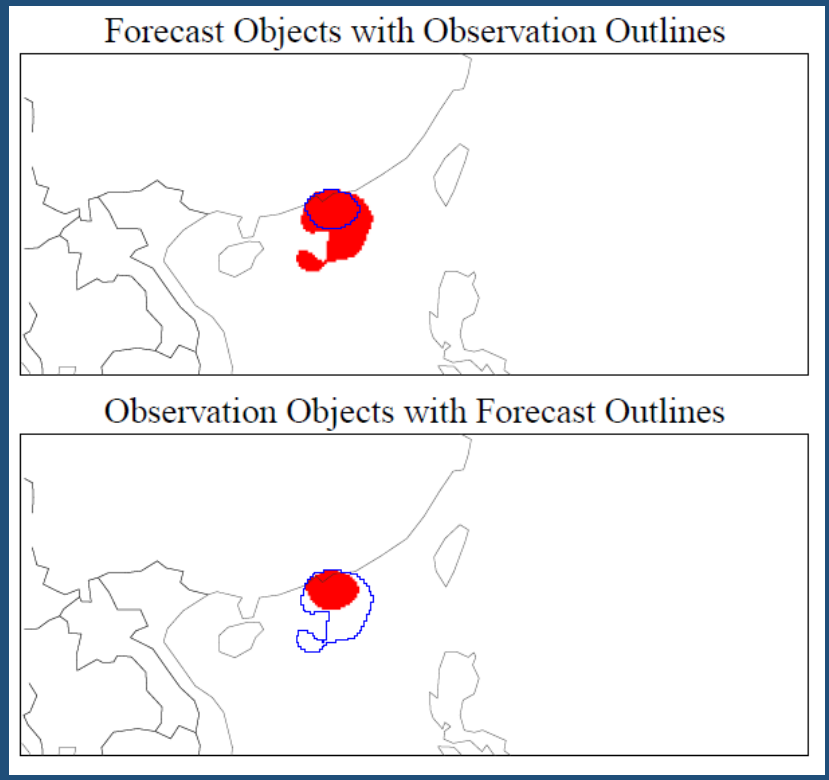
SuperT Hato / ***Gale Wind*** *or above*



	Forecast	Observation
Model	TrialModel	
Field	wind_speed	wind_speed
Level	*,*	*,*
Units	m/s	m/s
Initial	1970 01 01 00:00:00	1970 01 01 00:00:00
Valid	1970 01 01 00:00:00	1970 01 01 00:00:00
Accum	00:00:00	00:00:00

Centroid/Boundary	2.00	4.00
Convex Hull/Angle	0.00	1.00
Area/Intersection Area	1.00	2.00
Complexity/Intensity	0.00	0.00
Total Interest Thresh	0.70	

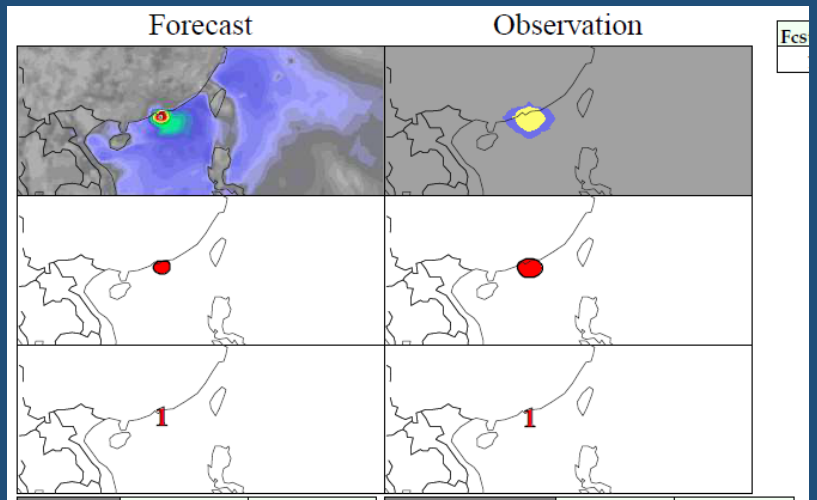
	Forecast	Observation
Mask M/G/P	on/off/off	on/off/off
Raw Thresh	>=0.0	>=0.0
Conv Radius	5	5
Conv Thresh	>=17.5	>=17.5
Area Thresh	>=0.0	>=0.0
Inten Thresh	p100>=0.0	p100>=0.0
Merge Thresh	>=16.5	>=16.5
Merging	thresh	thresh
Matching	match/merge	
Simple/M/U	1/1/0	1/1/0
Area	715	293
Area M/U	715/0	293/0
Cluster	1	1
MMI	0.8712	0.8712
MMI (F+O)	0.8712	



NCEP GFS

Base Time: 2017-08-21 12Z, 36-hour f/c
Valid Time: 2017-08-23 00Z

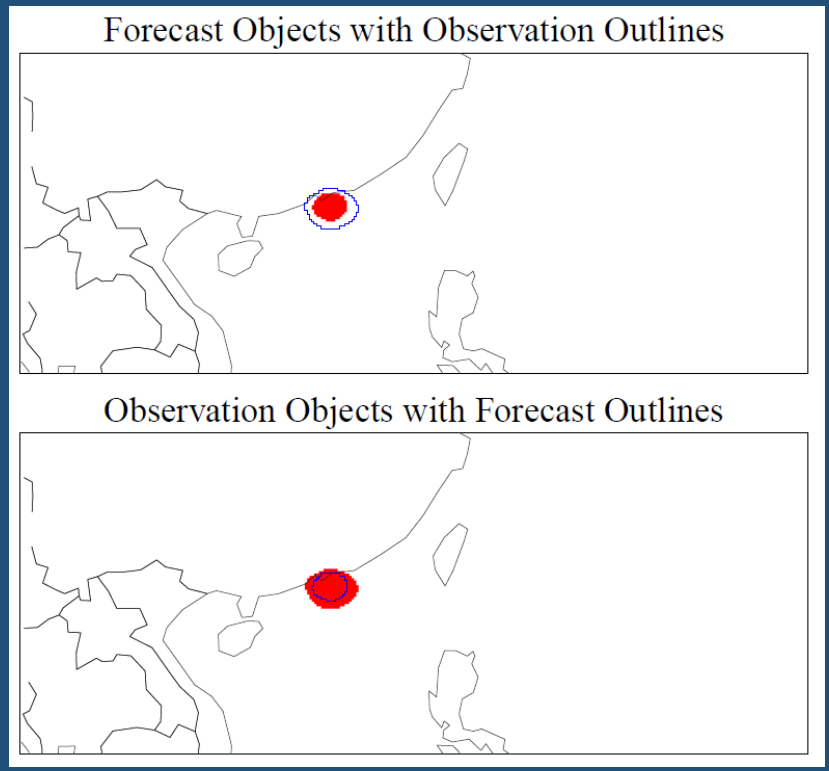
SuperT Hato / ***Gale Wind*** *or above*



	Forecast	Observation
Model	TrialModel	
Field	wind_speed	wind_speed
Level	*,*	*,*
Units	m/s	m/s
Initial	1970 01 01 00:00:00	1970 01 01 00:00:00
Valid	1970 01 01 00:00:00	1970 01 01 00:00:00
Accum	00:00:00	00:00:00

Centroid/Boundary	2.00	4.00
Convex Hull/Angle	0.00	1.00
Area/Intersection Area	1.00	2.00
Complexity/Intensity	0.00	0.00
Total Interest Thresh	0.70	

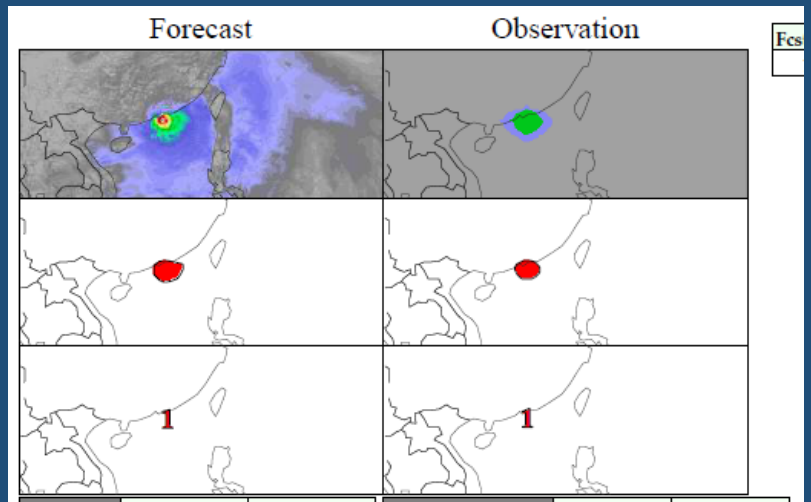
	Forecast	Observation
Mask M/G/P	on/off/off	on/off/off
Raw Thresh	>=0.0	>=0.0
Conv Radius	5	5
Conv Thresh	>=17.5	>=17.5
Area Thresh	>=0.0	>=0.0
Inten Thresh	p100>=0.0	p100>=0.0
Merge Thresh	>=16.5	>=16.5
Merging	thresh	thresh
Matching	match/merge	
Simple/M/U	1/1/0	1/1/0
Area	139	293
Area M/U	139/0	293/0
Cluster	1	1
MMI	0.9509	0.9509
MMI (F+O)	0.9509	



JMA GSM

Base Time: 2017-08-21 12Z, 36-hour f/c
Valid Time: 2017-08-23 00Z

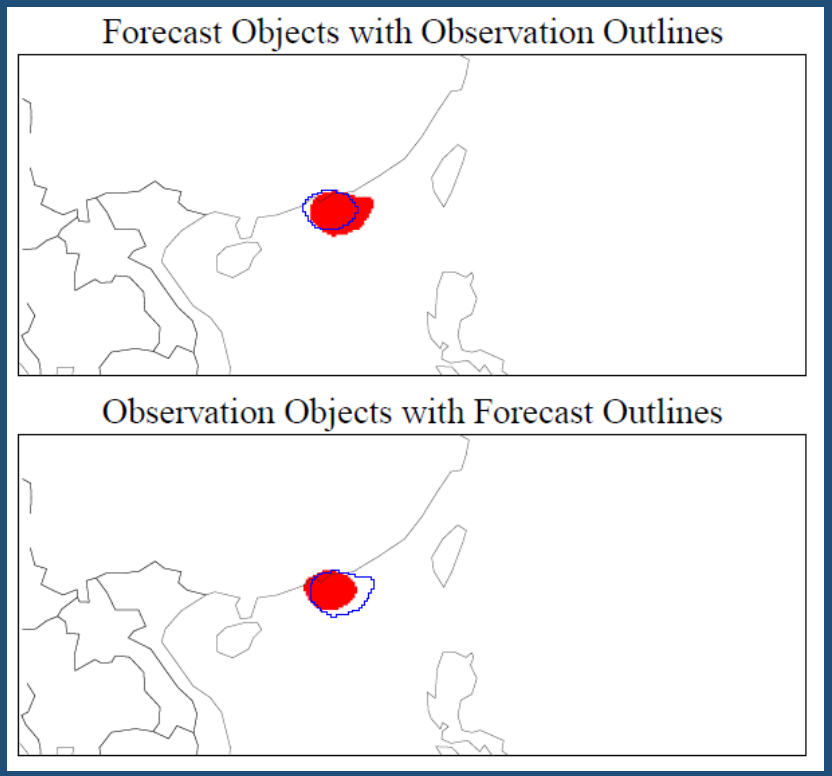
SuperT Hato / ***Gale Wind*** *or above*



	Forecast	Observation
Model	TrialModel	
Field	wind_speed	wind_speed
Level	*,*	*,*
Units	m/s	m/s
Initial	1970 01 01 00:00:00	1970 01 01 00:00:00
Valid	1970 01 01 00:00:00	1970 01 01 00:00:00
Accum	00:00:00	00:00:00

Centroid/Boundary	2.00	4.00
Convex Hull/Angle	0.00	1.00
Area/Intersection Area	1.00	2.00
Complexity/Intensity	0.00	0.00
Total Interest Thresh	0.70	

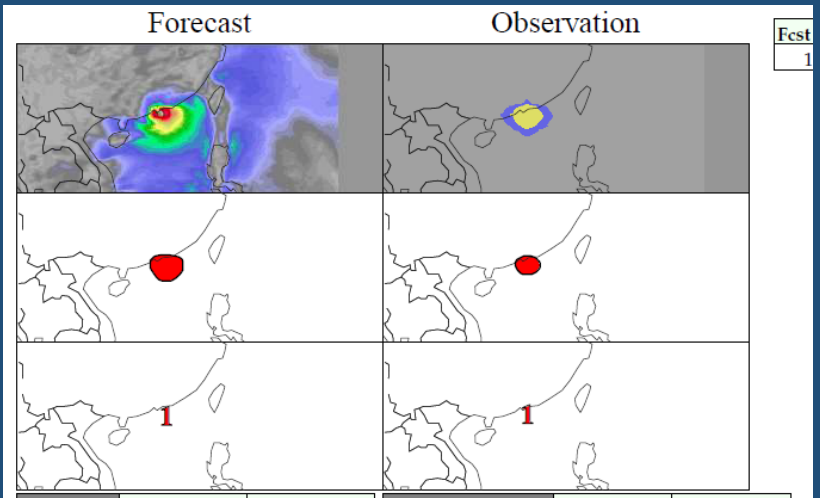
	Forecast	Observation
Mask M/G/P	on/off/off	on/off/off
Raw Thresh	>=0.0	>=0.0
Conv Radius	5	5
Conv Thresh	>=17.5	>=17.5
Area Thresh	>=0.0	>=0.0
Inten Thresh	p100>=0.0	p100>=0.0
Merge Thresh	>=16.5	>=16.5
Merging	thresh	thresh
Matching	match/merge	
Simple/M/U	1/1/0	1/1/0
Area	389	293
Area M/U	389/0	293/0
Cluster	1	1
MMI	0.9564	0.9564
MMI (F+O)	0.9564	



ECMWF IFS

Base Time: 2017-08-21 12Z, 36-hour f/c
Valid Time: 2017-08-23 00Z

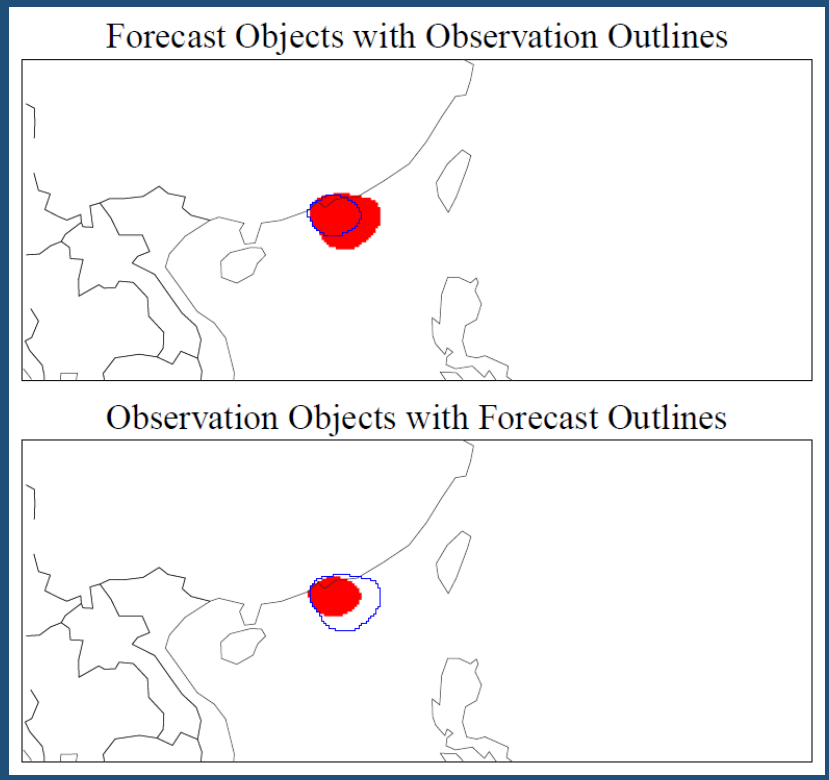
SuperT Hato / ***Gale Wind*** *or above*



	Forecast	Observation
Model	TrialModel	
Field	wind_speed	wind_speed
Level	*,*	*,*
Units	m/s	m/s
Initial	1970 01 01 00:00:00	1970 01 01 00:00:00
Valid	1970 01 01 00:00:00	1970 01 01 00:00:00
Accum	00:00:00	00:00:00

Centroid/Boundary	2.00	4.00
Convex Hull/Angle	0.00	1.00
Area/Intersection Area	1.00	2.00
Complexity/Intensity	0.00	0.00
Total Interest Thresh	0.70	

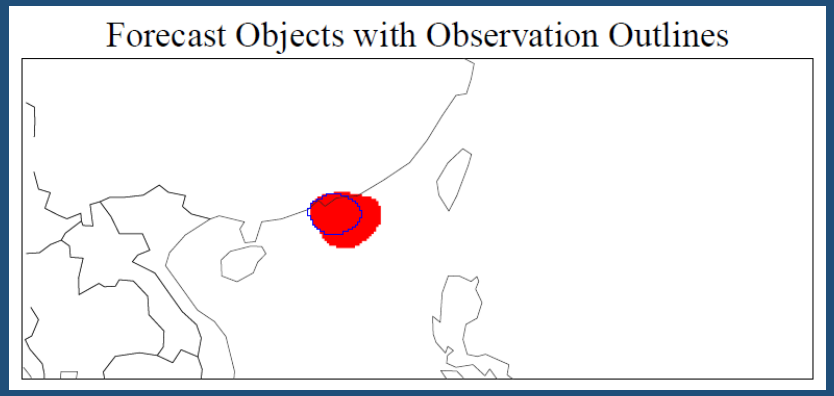
	Forecast	Observation
Mask M/G/P	on/off/off	on/off/off
Raw Thresh	>=0.0	>=0.0
Conv Radius	5	5
Conv Thresh	>=17.5	>=17.5
Area Thresh	>=0.0	>=0.0
Inten Thresh	p100>=0.0	p100>=0.0
Merge Thresh	>=16.5	>=16.5
Merging	thresh	thresh
Matching	match/merge	
Simple/M/U	1/1/0	1/1/0
Area	562	293
Area M/U	562/0	293/0
Cluster	1	1
MMI	0.9268	0.9268
MMI (F+O)	0.9268	

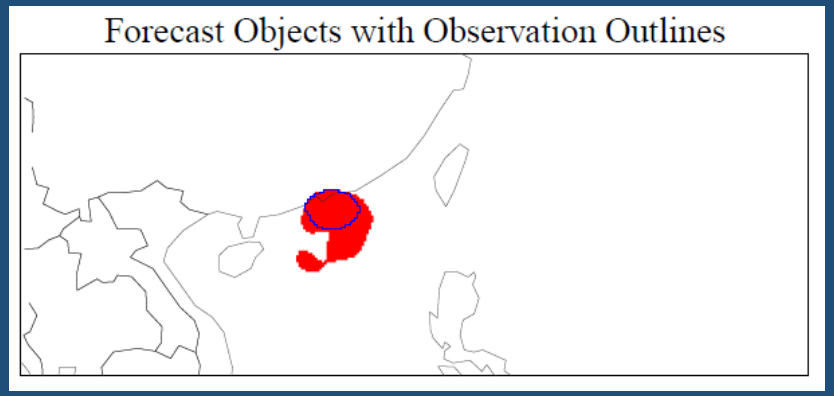


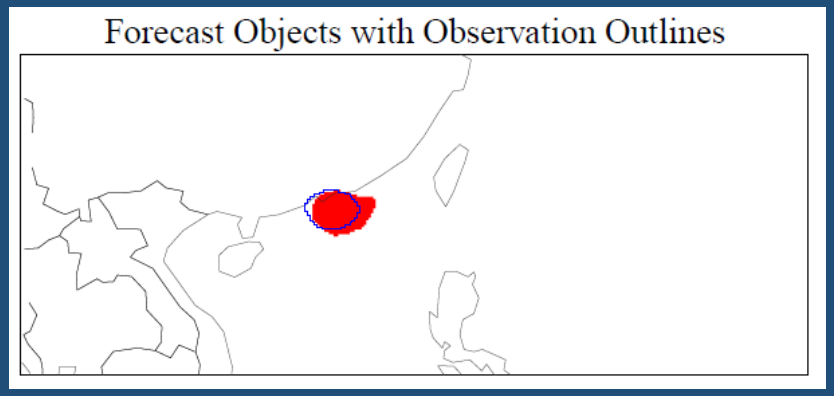
UKMO UM

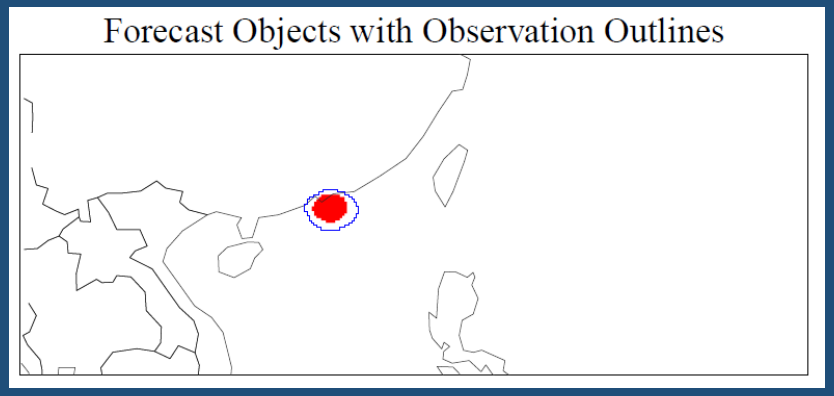
SuperT Hato 36-hour forecast Based at 2017/08/21 12Z Gale

WHICH IS BETTER ?



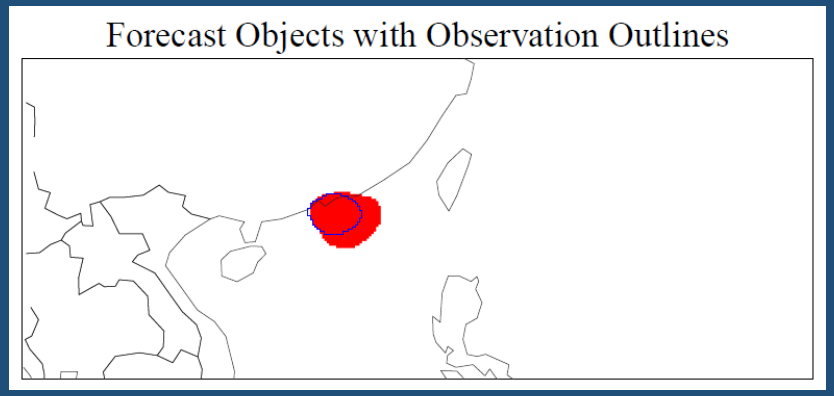




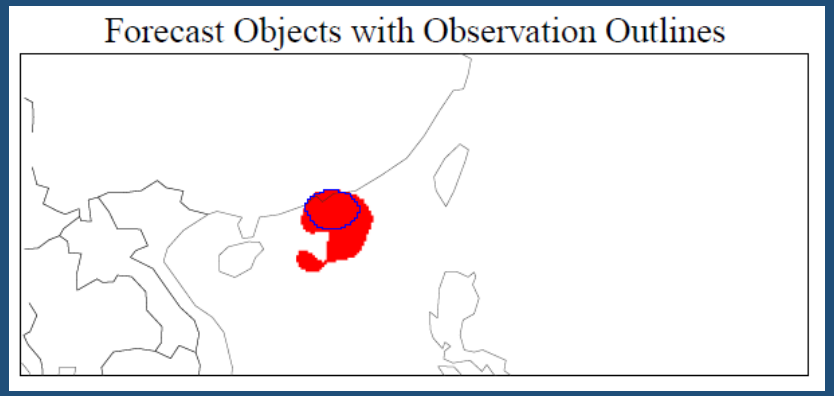


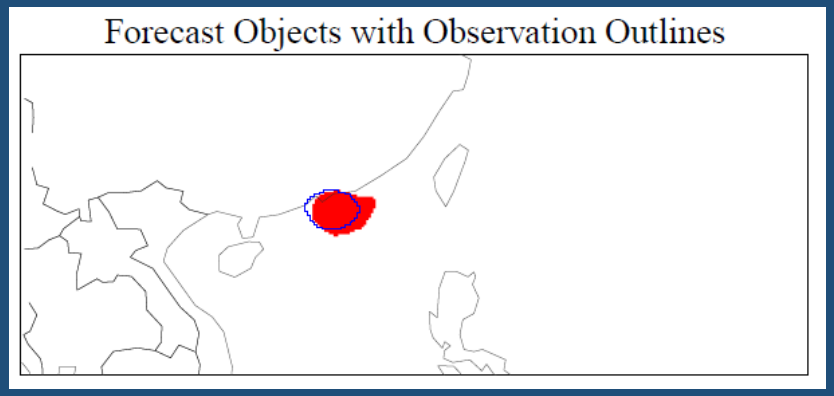
SuperT Hato 36-hour forecast Based at 2017/08/21 12Z Gale

UKMO: MMI 0.93

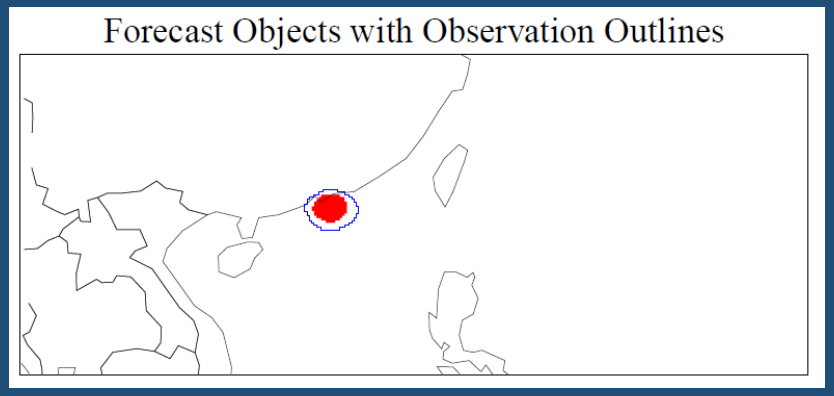


NCEP: MMI 0.87





ECMWF: MMI 0.96



JMA: MMI 0.95

Experiment

- Goal:
  - See whether performance of models as ranked by MODE MMI are consistent with subjective rankings by **forecasters**.
- Setup:
  - Without being shown the above object-based verification, 9 forecasters were asked to rank the performance of the TC structure by different models subjectively.
- Result …

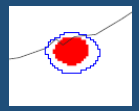
Subjective Ranking of Hato:

	Best	Better	Worse	Worst
Forecaster #1	UKMO	ECMWF	NCEP	JMA
Forecaster #2	ECMWF	UKMO	NCEP	JMA
Forecaster #3	ECMWF	JMA	UKMO	NCEP
Forecaster #4	ECMWF	JMA	NCEP	UKMO
Forecaster #5	ECMWF	UKMO	JMA	NCEP
Forecaster #6	ECMWF	UKMO	JMA	NCEP
Forecaster #7	ECMWF	JMA	UKMO	NCEP
Forecaster #8	ECMWF	UKMO	JMA	NCEP
Forecaster #9	UKMO	ECMWF	NCEP	JMA
MODE MMI	**ECMWF (0.96)**	**JMA (0.95)**	**UKMO (0.93)**	**NCEP (0.87)**

Discussions

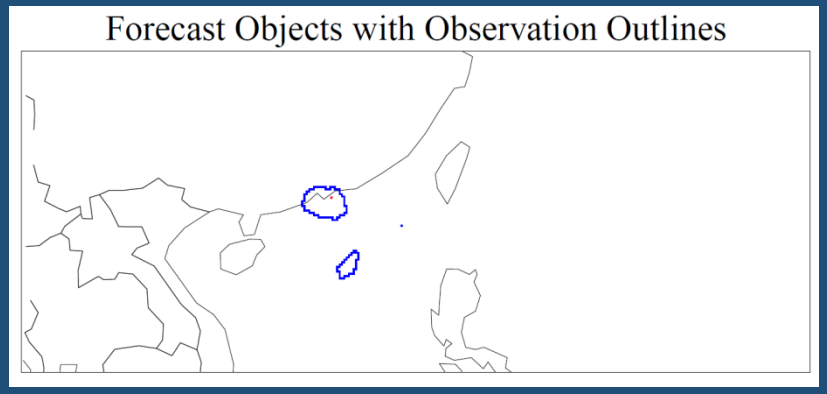
- Most forecasters ranked ECMWF as the best and half of them ranked NCEP as the worst, generally in line with MODE MMI ranking.

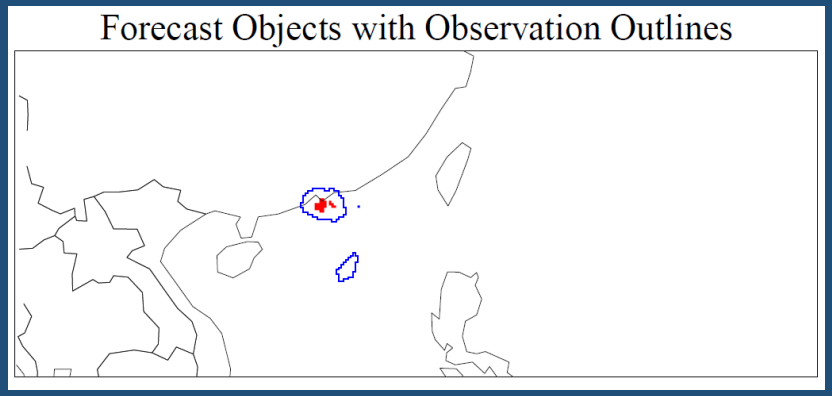
- Many forecasters ranked JMA poorer than its MODE MMI rank (second best), probably as a result of large area of misses, which subjectively are “wronger”.

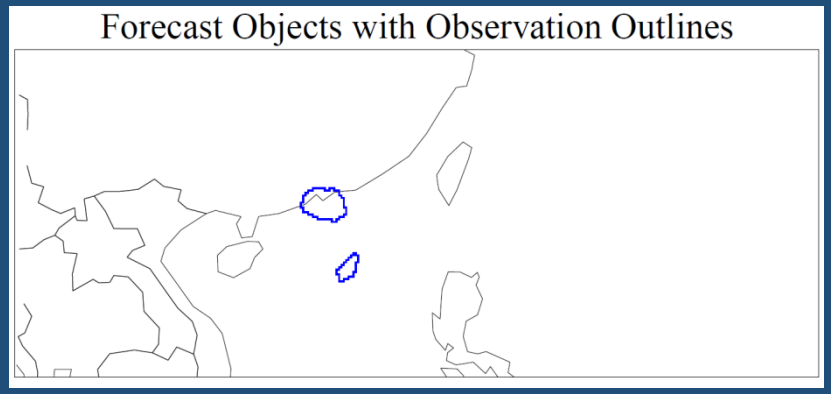


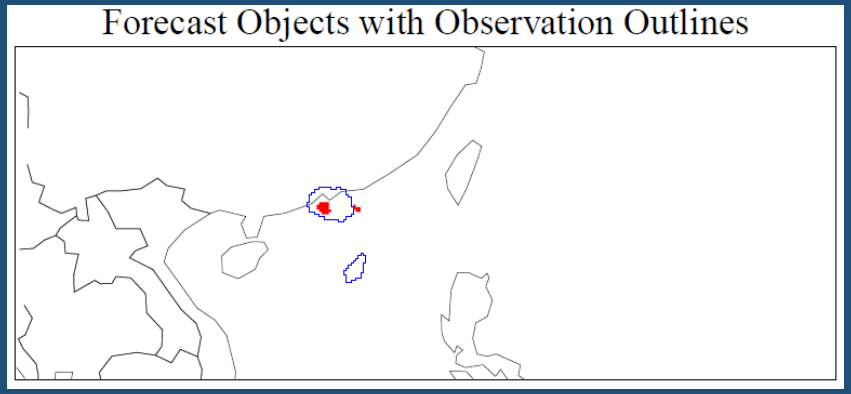
- Different weighting to attributes or fine-tune the calibration curves are to be designed

STS Pakhar 24-hour forecast Based at 2017/08/26 00Z Gale



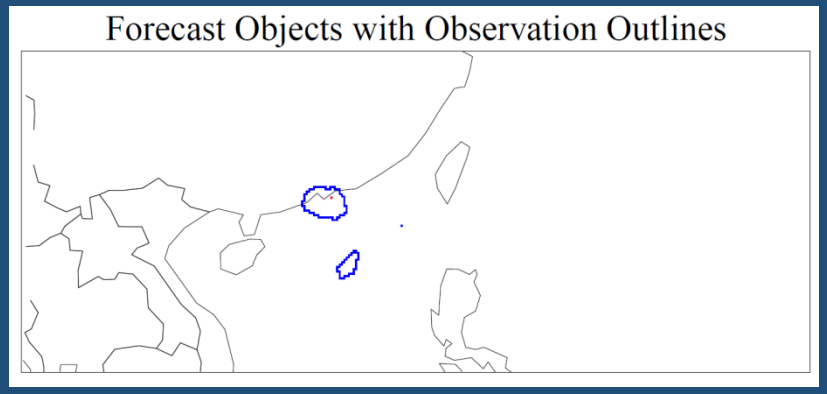




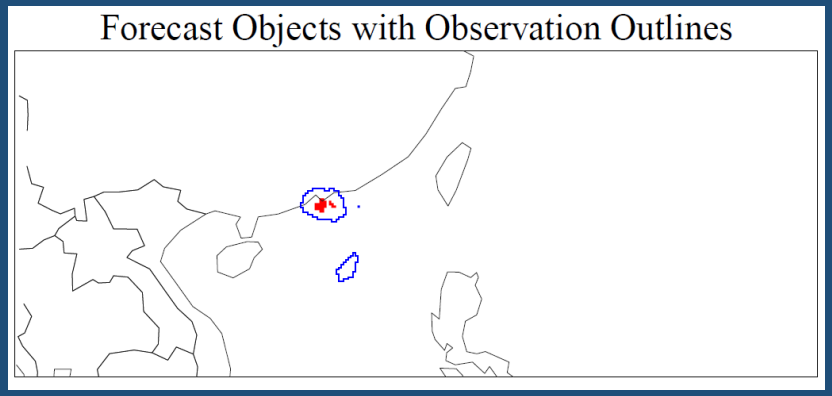


STS Pakhar 24-hour forecast Based at 2017/08/26 00Z Gale

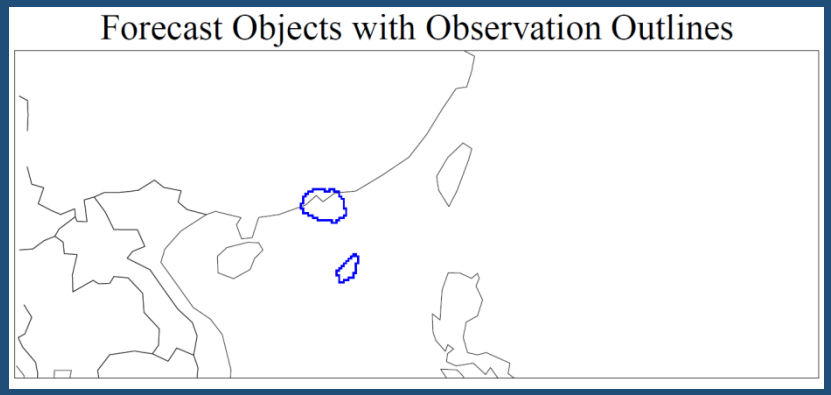
NCEP: MMI 0.43



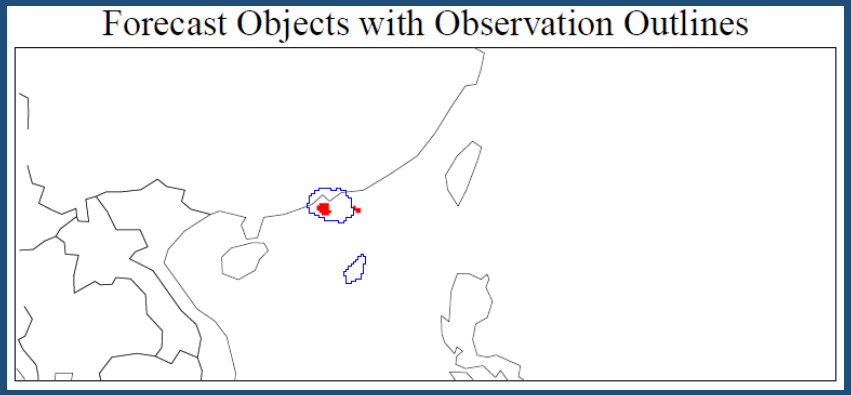
ECMWF: MMI 0.85



JMA: MMI 0.00



UKMO: MMI 0.74



Subjective Ranking of Pakhar

	Best	Better	Worse	Worst
Forecaster #1	ECMWF	UKMO	NCEP	JMA
Forecaster #2	ECMWF	UKMO	NCEP	JMA
Forecaster #3	ECMWF	UKMO	NCEP	JMA
Forecaster #4	ECMWF	UKMO	NCEP	JMA
Forecaster #5	ECMWF	UKMO	NCEP	JMA
Forecaster #6	ECMWF	UKMO	NCEP	JMA
Forecaster #7	ECMWF	UKMO	NCEP	JMA
Forecaster #8	ECMWF	UKMO	NCEP	JMA
Forecaster #9	ECMWF	UKMO	NCEP	JMA
MODE MMI	**ECMWF**	**UKMO**	**NCEP**	**JMA**

Discussion

- MODE MMI ranking matches perfectly with forecasters' subjective ranking in this case.

- Likely due to significant difference in performance and clear-cut false alarms

香港天文台
HONG KONG OBSERVATORY

Comments on performance of TC structure by NWPs

- ECMWF's 36-hour forecast captured the gale wind structures of SuperT Hato quite well, revealing the usefulness of dynamical model guidance in this case.

- Most global NWP's 24-hour forecast under-estimated gale winds of STS Pakhar, however.

香港天文台
HONG KONG OBSERVATORY

Conclusions

- To support impact-based warnings, verification of TC wind structures are necessary.
- MODE provides an object-based verification method. MODE MMI seems to be useful for ranking forecast performance of TC wind structures.
- Tuning of the scoring curves and weightings to be explored to better reflect "what" forecasters or users are concerned.

香港天文台
HONG KONG OBSERVATORY

Thank you very much