



Verification Against Multiple Analyses

Forecast

Own Analysis

Others Analysis

Analysis

	JMA	ECMWF	UKMO	NCEP	DWD	KMA	CMC	CMA	BOM
JMA	Own Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis
ECMWF	Others Analysis	Own Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis
UKMO	Others Analysis	Others Analysis	Own Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis
NCEP	Others Analysis	Others Analysis	Others Analysis	Own Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis
DWD	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Own Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis
KMA	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Own Analysis	Others Analysis	Others Analysis	Others Analysis
CMC	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Own Analysis	Others Analysis	Others Analysis
CMA	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Own Analysis	Others Analysis
BOM	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Others Analysis	Own Analysis

Takafumi Kanehama

10-13 Mar. 2014, Melbourne

WGNE-29

Motivation

- Analysis fields are often used as reference fields (i.e. ‘truth’) for verification because they are easy to handle.
 - But they include analysis error, therefore not ‘truth’ value.
- The influence of analysis to scores can be estimated by **verifying forecasts from multiple centers against analyses from multiple centers.**
 - If the results are insensitive to analysis choice, the metric is reliable.
 - If not, it is difficult to draw any conclusion by just comparing own-analysis verification results.
 - Also, this may suggest analysis error is as large as forecast error.

Method

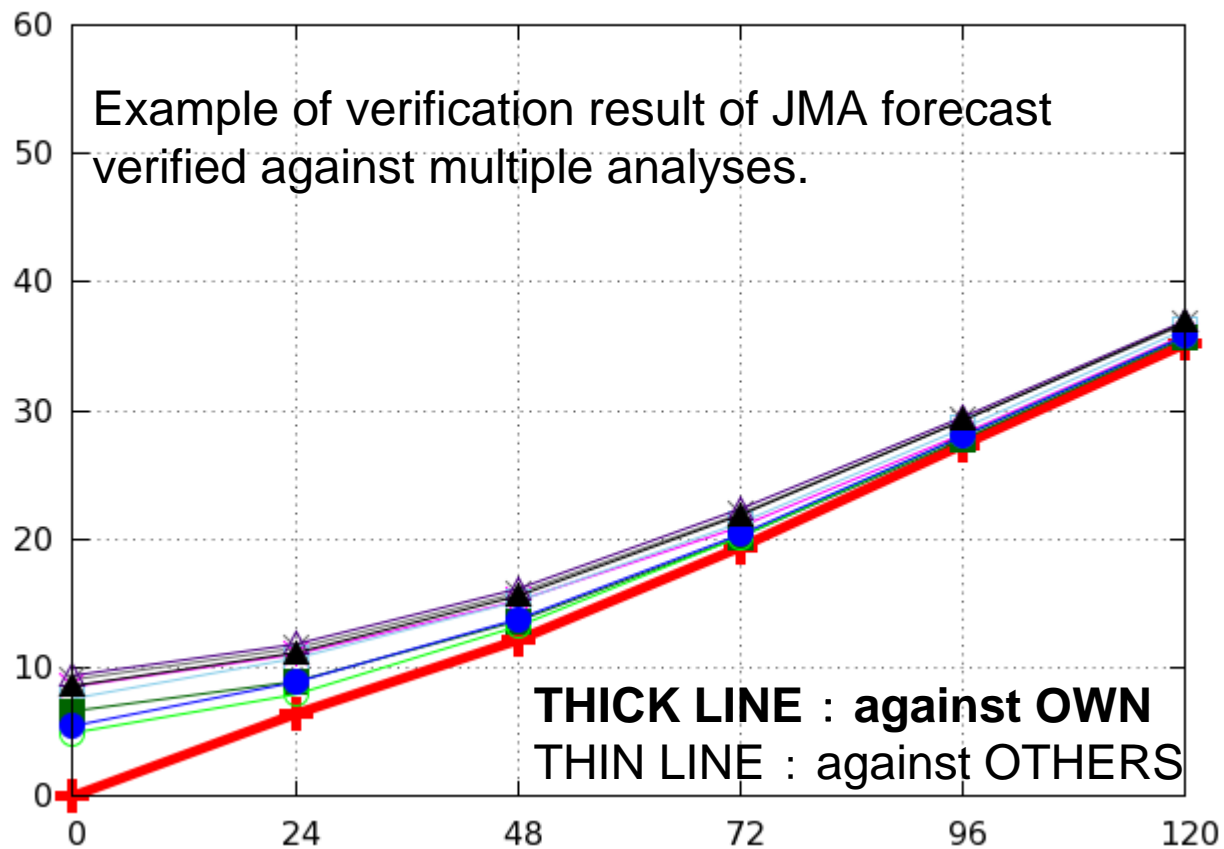
- Verification of multiple forecasts against multiple analyses
 - Using JMA DB :
 - 9 forecasts x 9 analyses : Z500, PSEA
 - 3 forecasts x 3 analyses : Z500, PSEA, T850, WS850 (or VW850)
- Period : 2012 July 01-30, 12UTC (30 samples)
 - 29 samples for verification using KMA data because of one missing data in DB
- Forecast Hour : T+120 hours
- Region : NH, TR, SH
- Score : RMSE, ACC, STDVE
- Climatology for ACC : ERA-Interim daily climatology (distributed by ECMWF as WMO LC-DNV)
- Resolution : interpolated to the coarsest resolution among forecast, analysis and climatology data
- Average : arithmetic average

9 X 9

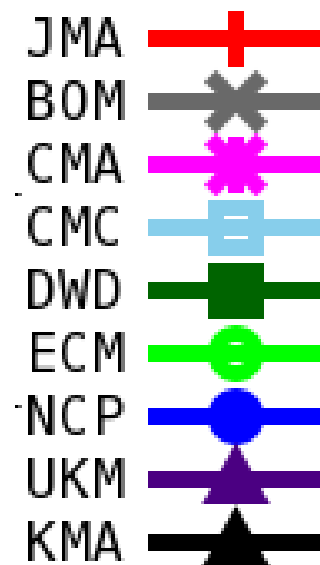
(JMA, BOM, CMA, CMC, DWD, ECM, NCP, UKM, KMA)

Z500, PSEA

fcst:JMA NHEM RMSE z500

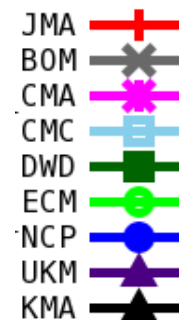
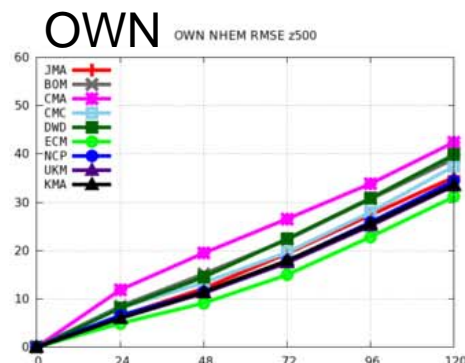
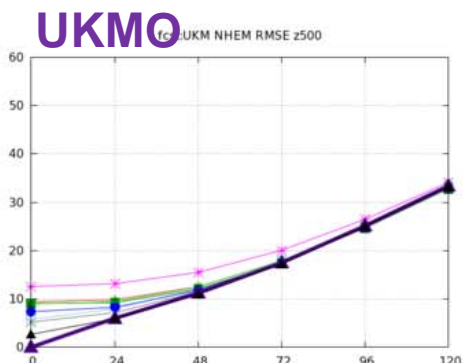
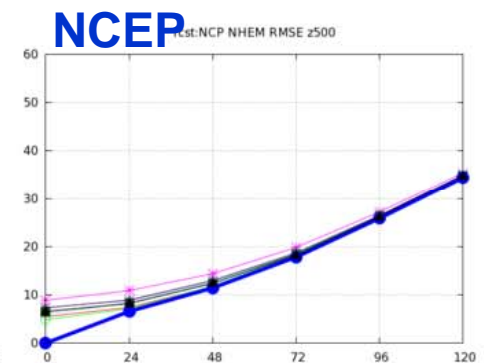
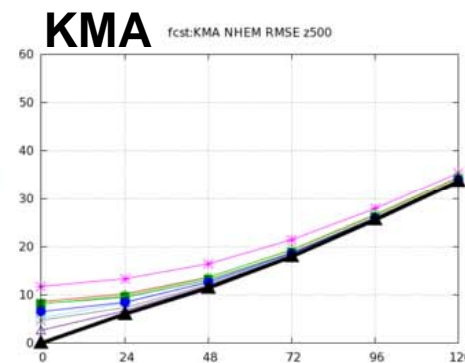
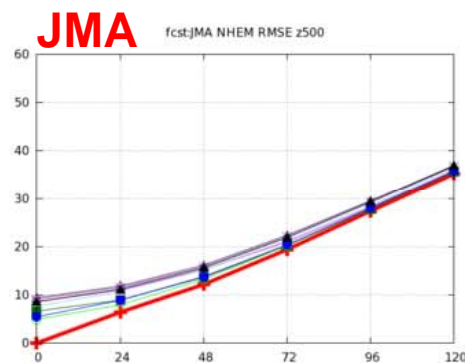
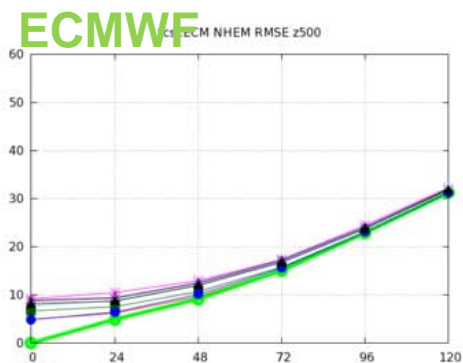
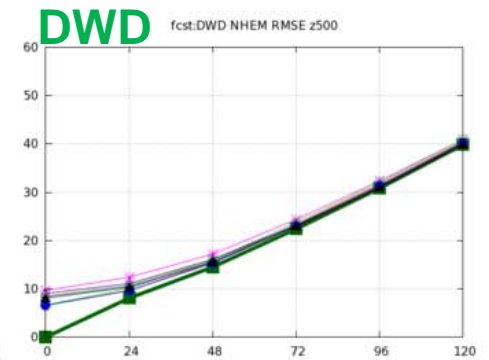
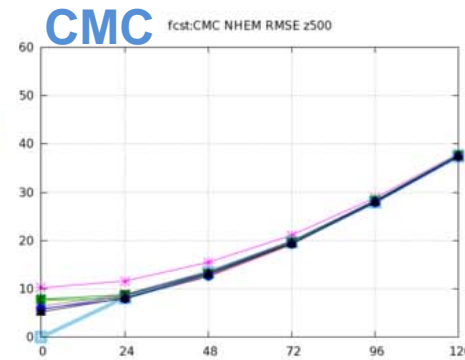
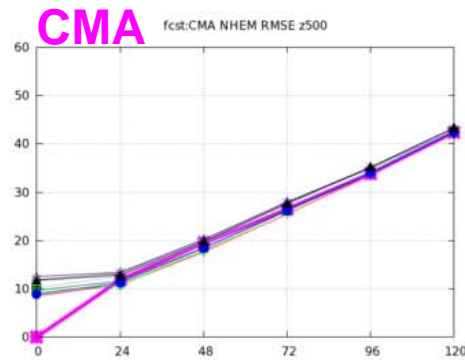
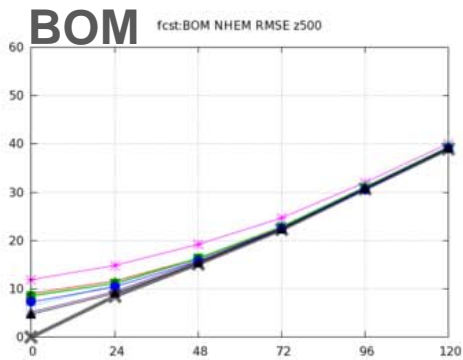


colored by centers



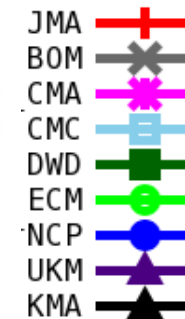
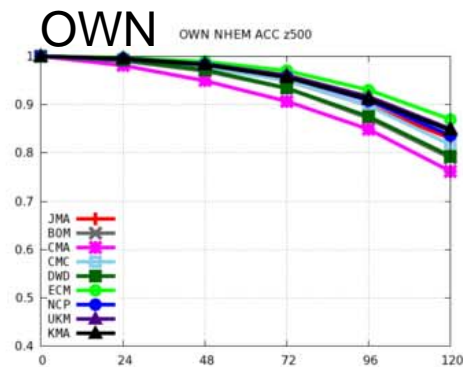
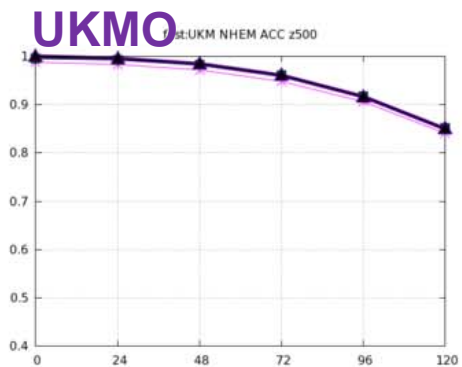
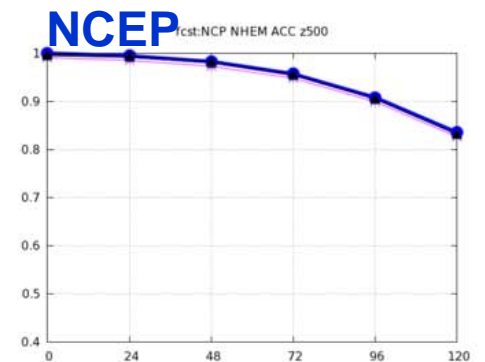
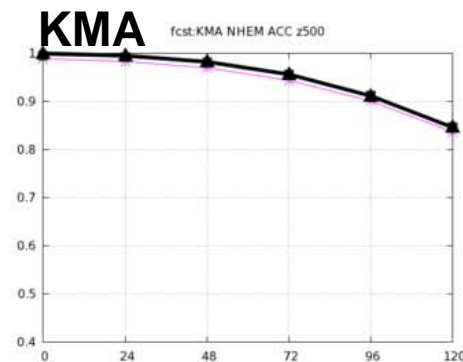
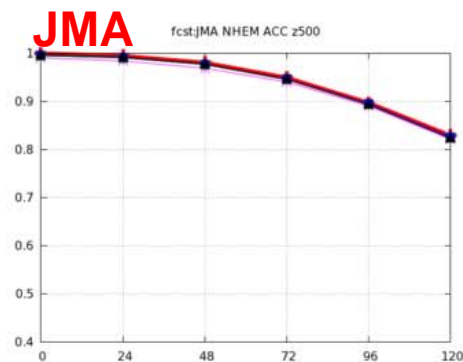
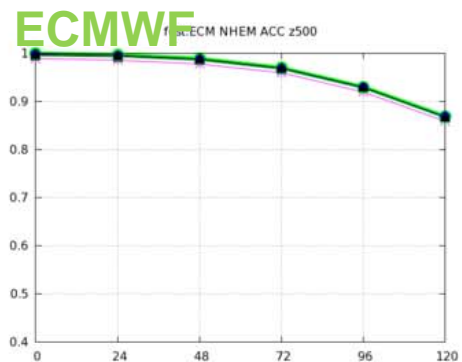
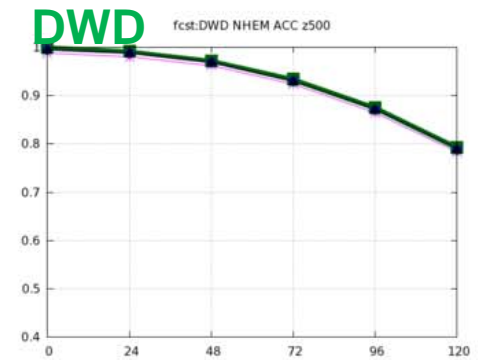
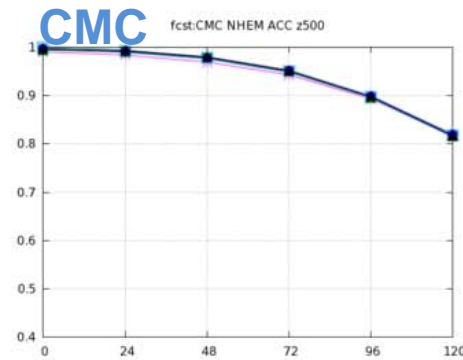
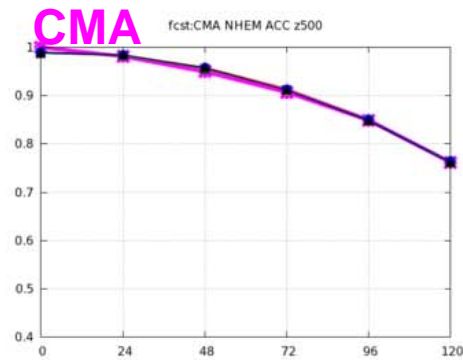
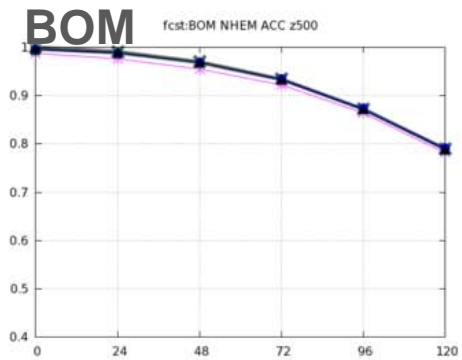


EXTRA-TROPICS



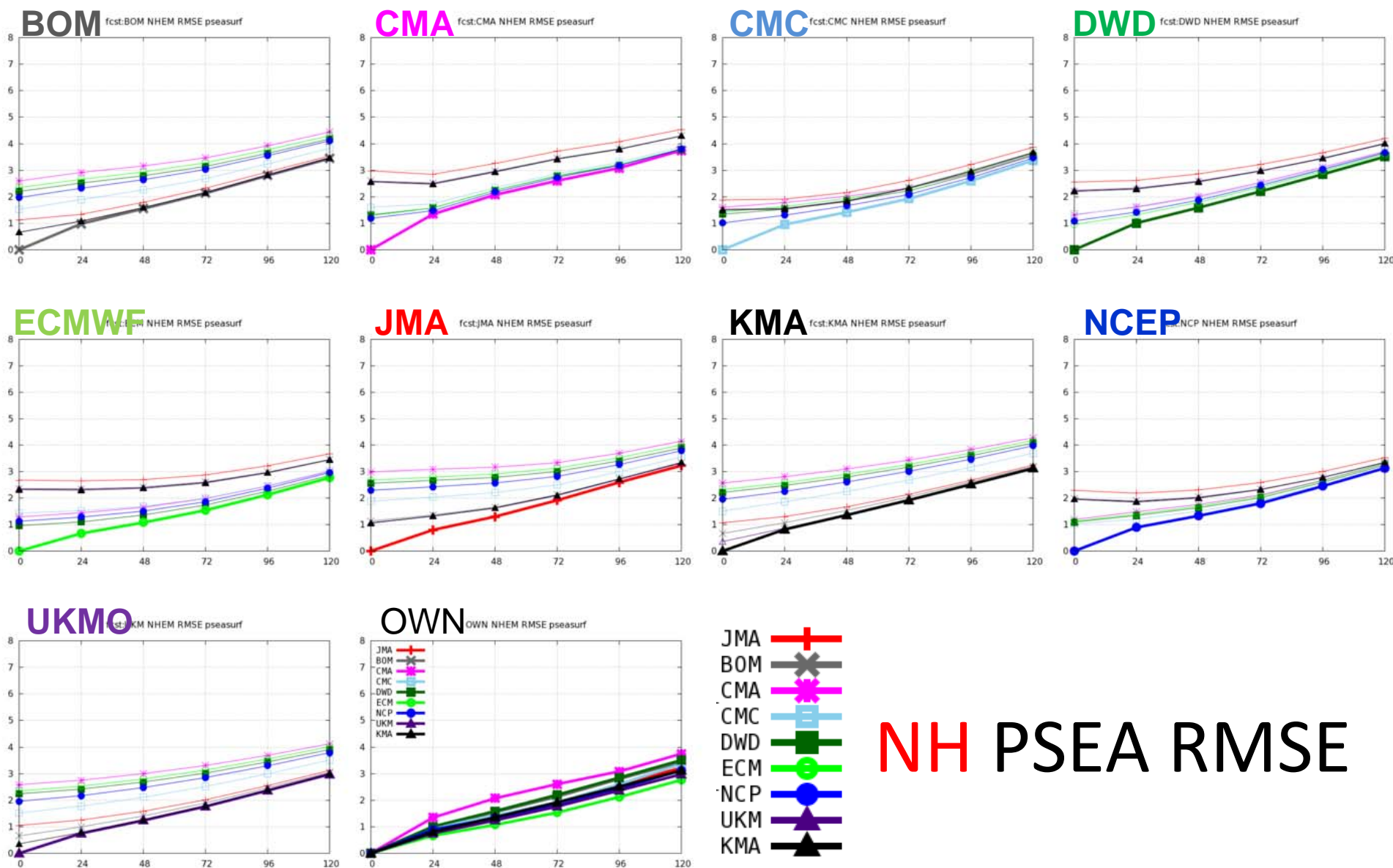
NH Z500 RMSE

- Up to day 2, the influences of analysis to score are relatively large compared to differences between OWN scores.
- Over day 2, the influences of analysis to score are small. Differences seem to be converged.
- The sensitivity differs center by center. Score variability of some centers converges faster than others.



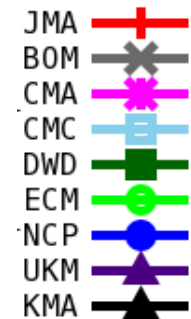
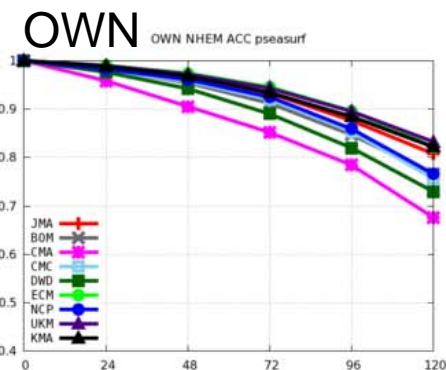
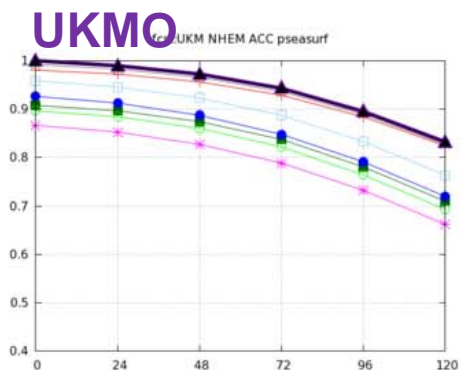
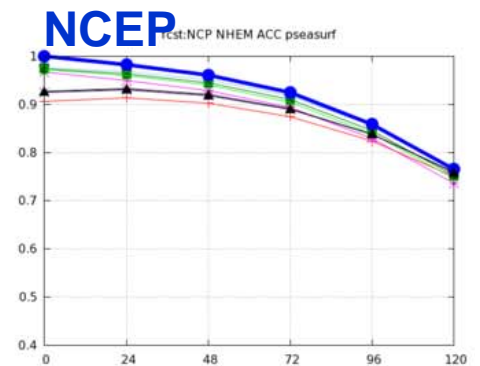
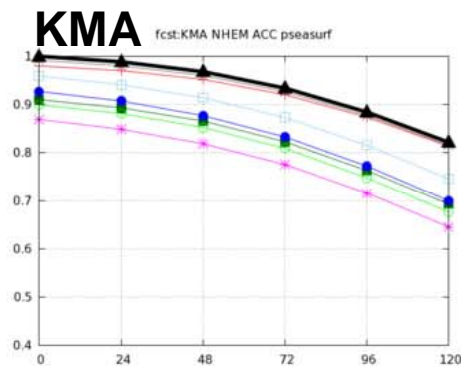
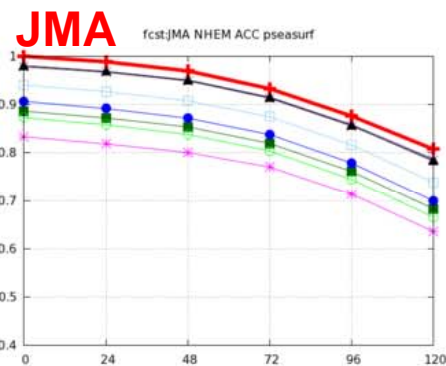
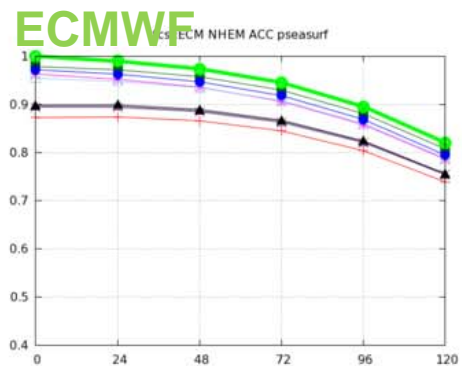
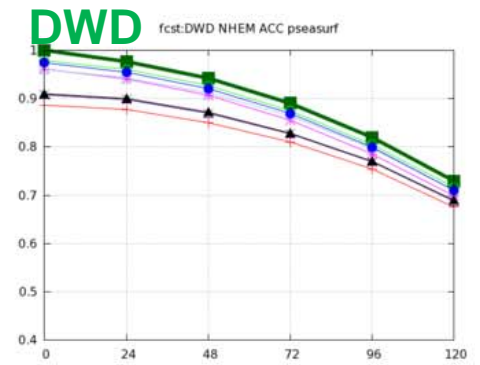
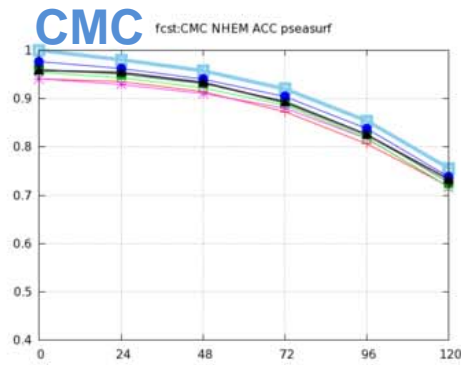
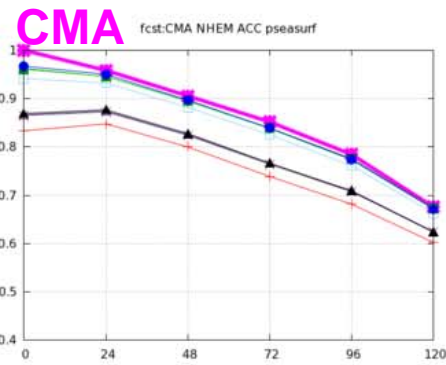
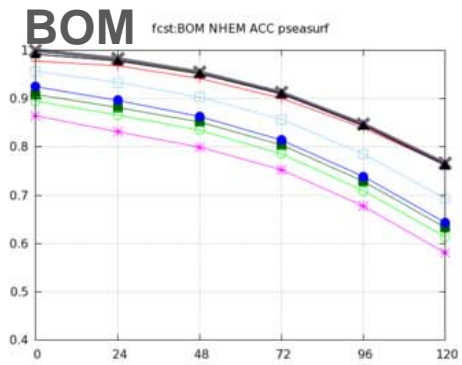
NH Z500 ACC

- ACC is not sensitive to choice of analysis through forecast time.



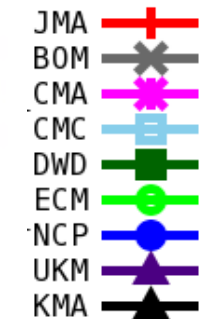
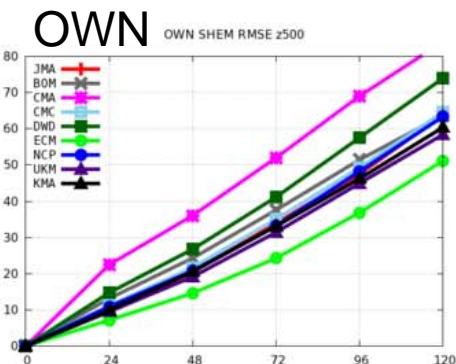
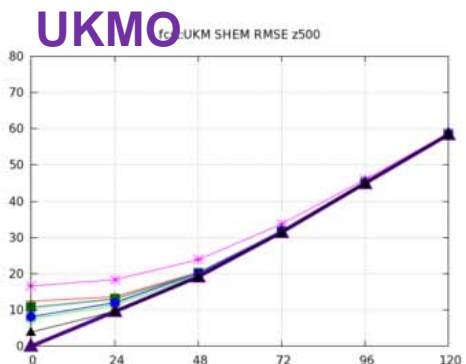
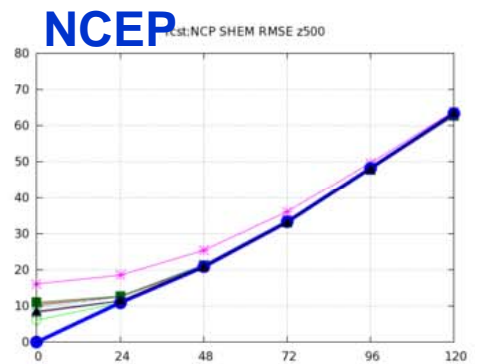
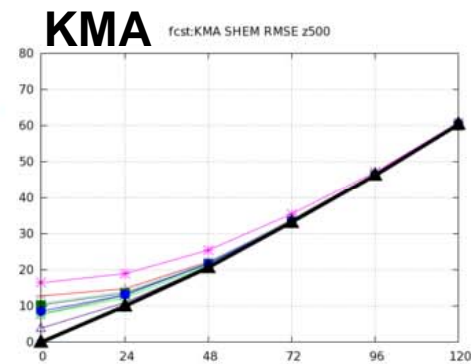
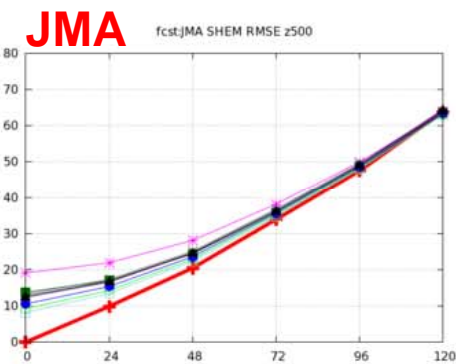
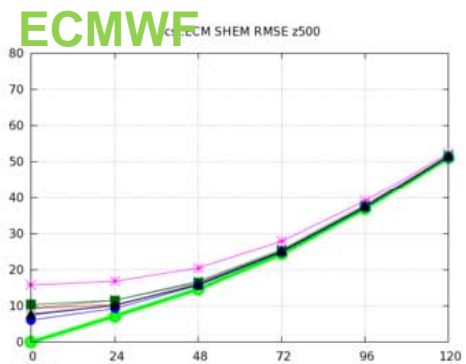
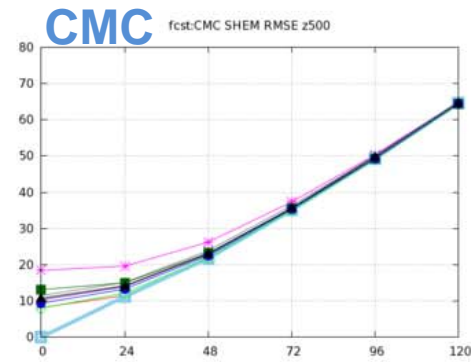
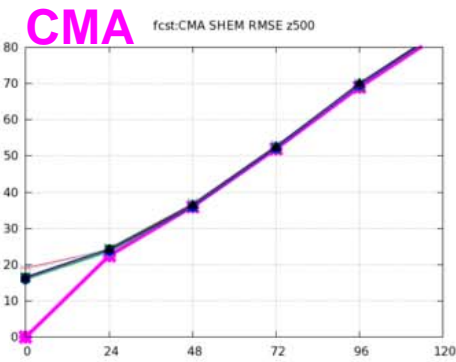
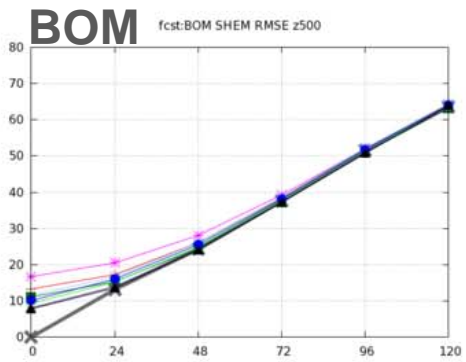
NH PSEA RMSE

- Score variability is not converged even at day 5 forecast and is larger than the difference of OWN scores.



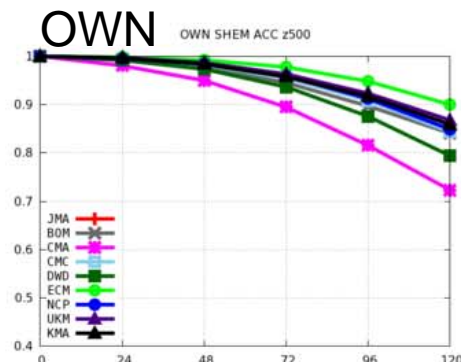
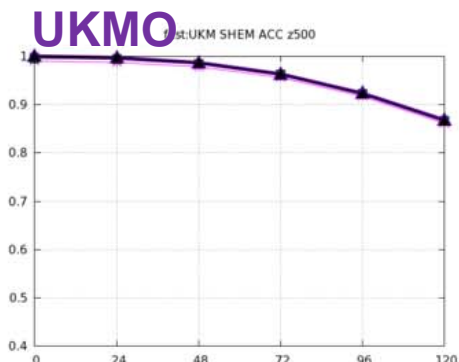
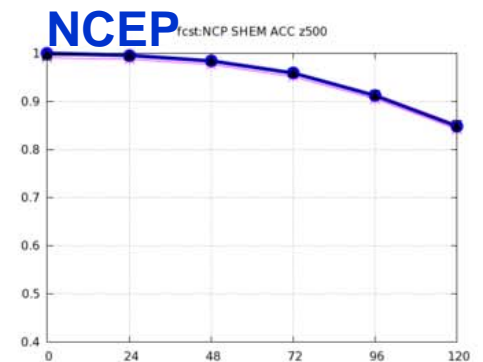
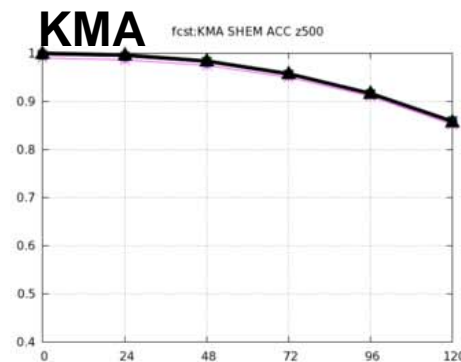
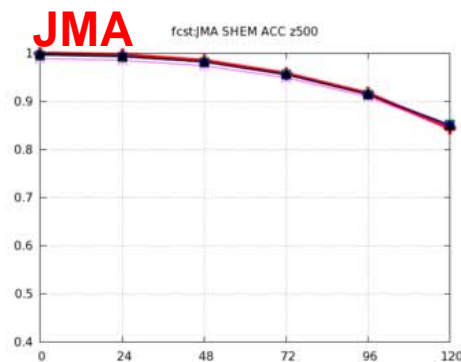
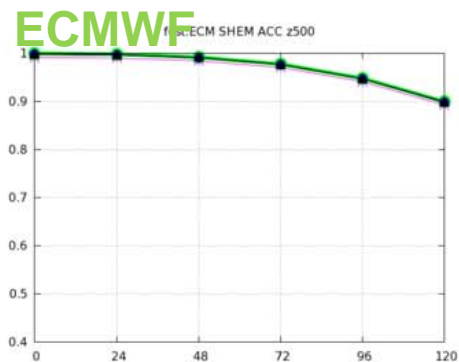
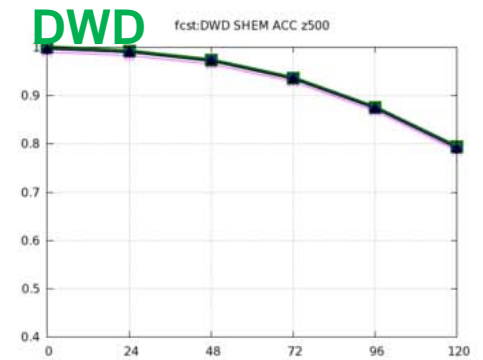
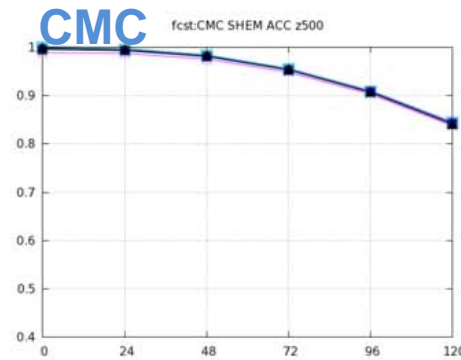
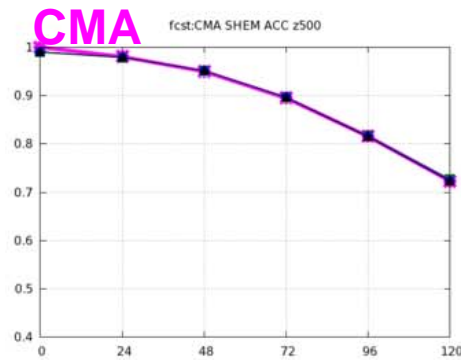
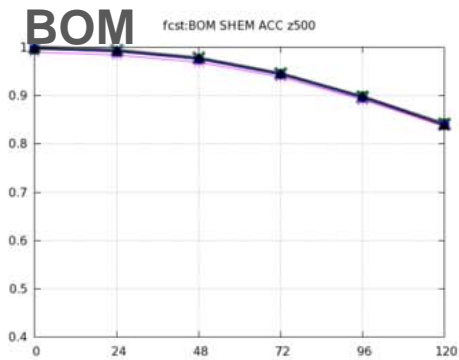
NH PSEA ACC

- Score variability is not converged even at day 5 forecast and is larger than the difference of OWN scores.
- Are there 'good' combinations such as (JMA,UKM,KMA,BOM) and (ECM,CMC,CMA,DWD,NCP)?



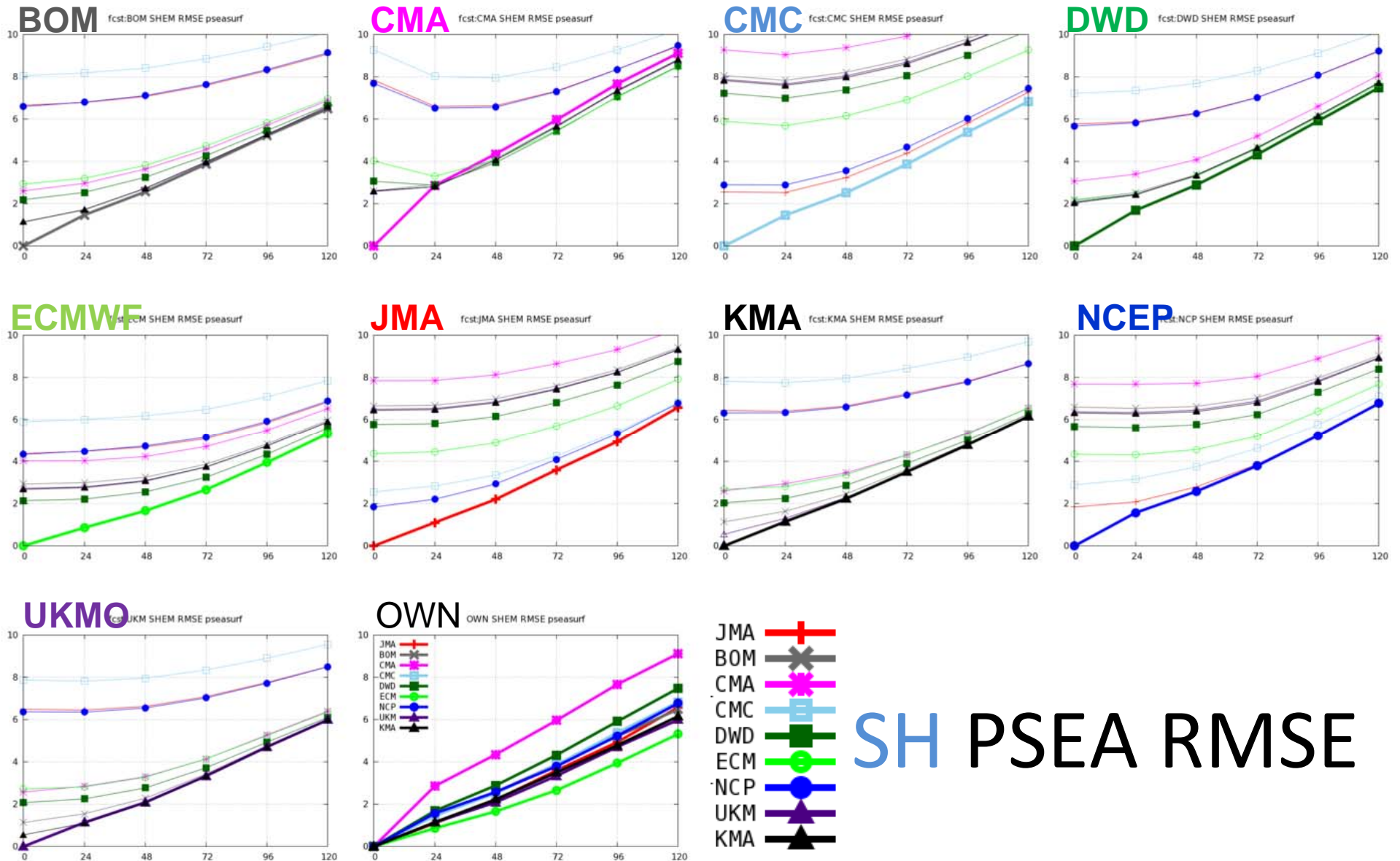
SH Z500 RMSE

- Up to day 2, the influences of analysis to score are relatively large compared to differences between OWN scores.
- Over day 2, the influences of analysis to score are small. Differences seem to be converged.
- The sensitivity differs center by center. Score variability of some centers converges faster than others.

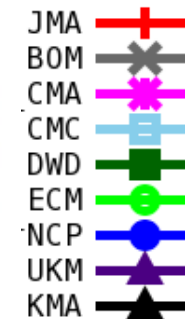
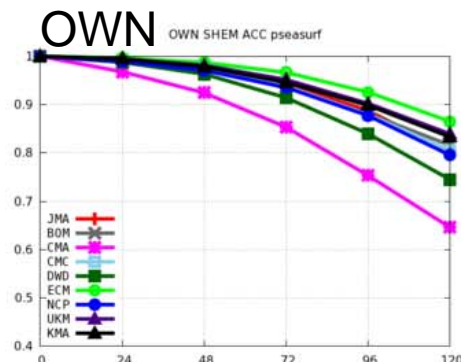
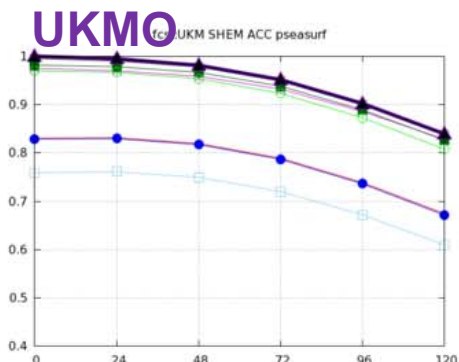
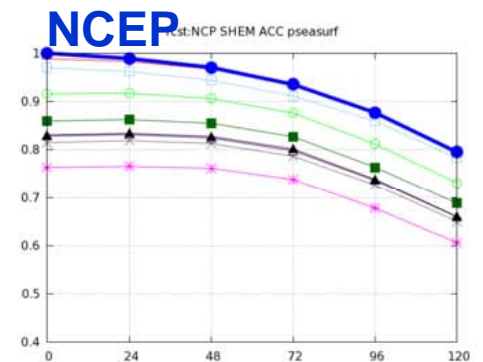
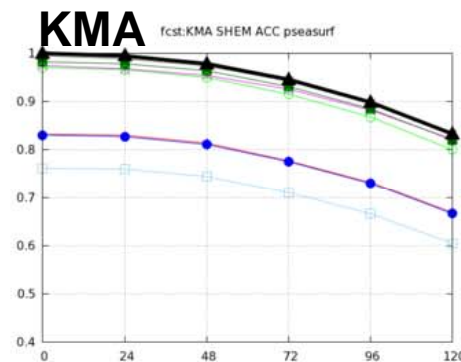
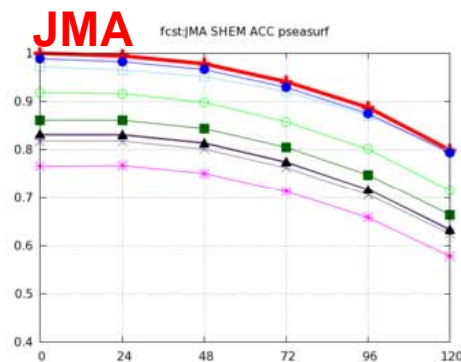
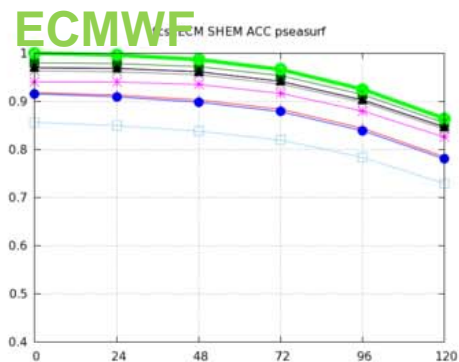
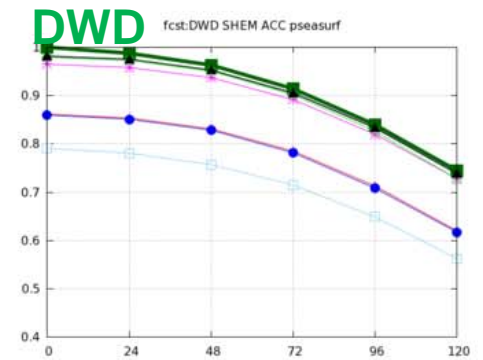
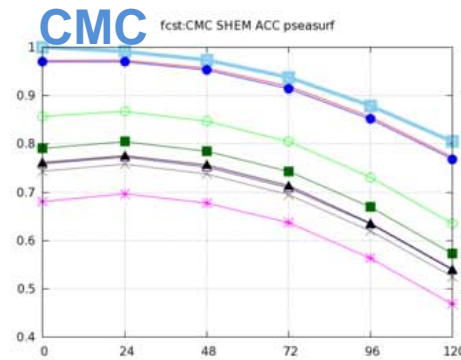
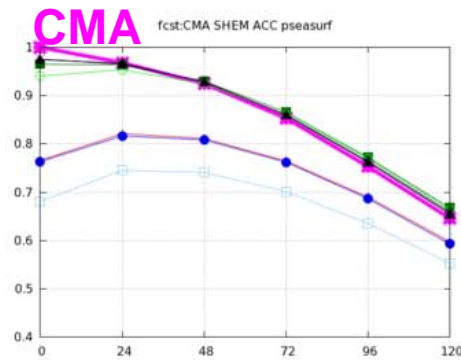
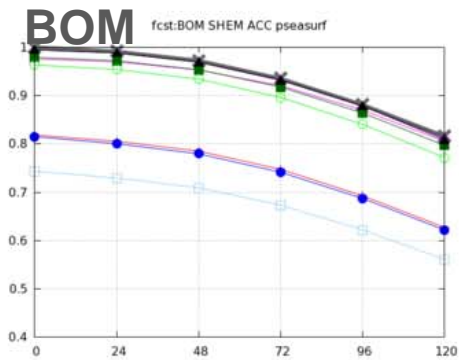


SH Z500 ACC

- ACC is not sensitive to choice of analysis through forecast time.



- Score variability is not converged even at day 5 forecast and is larger than the difference of OWN scores.

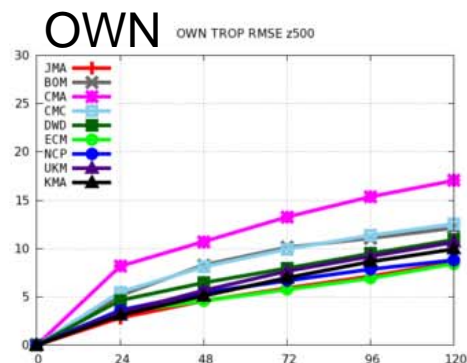
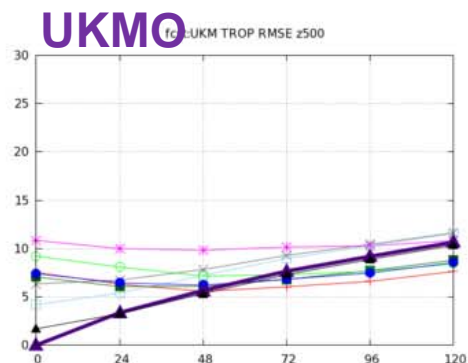
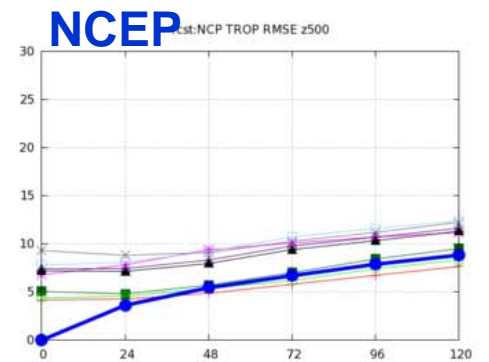
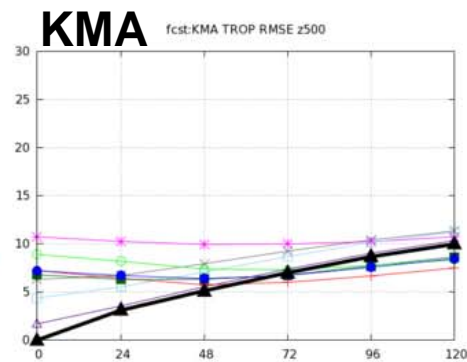
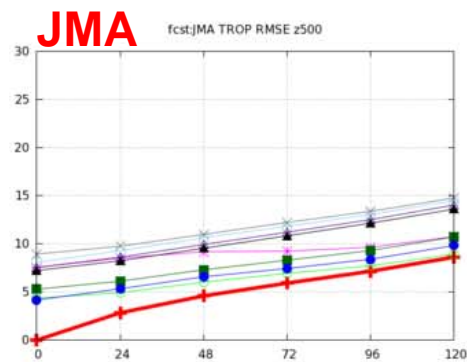
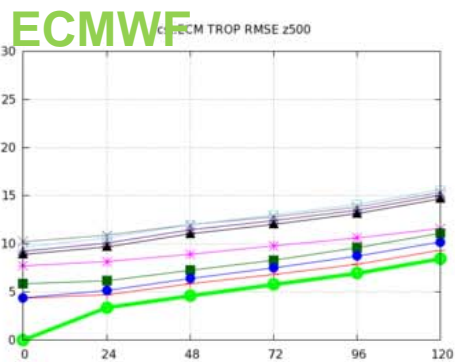
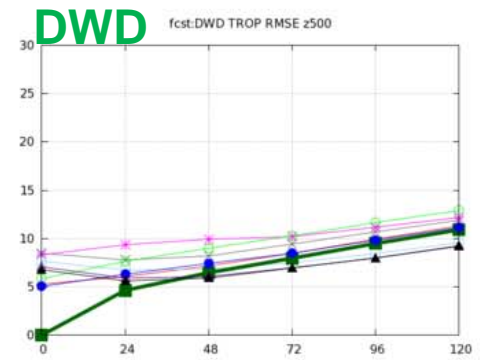
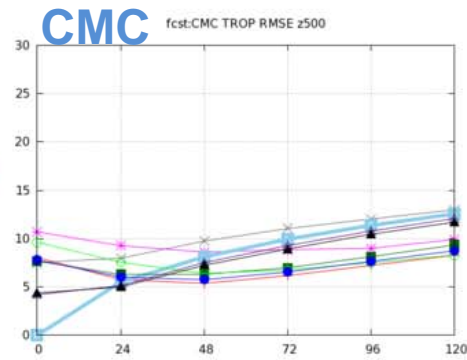
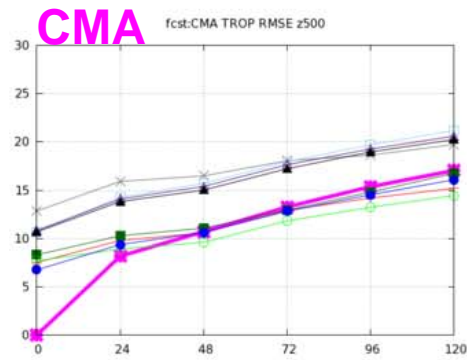
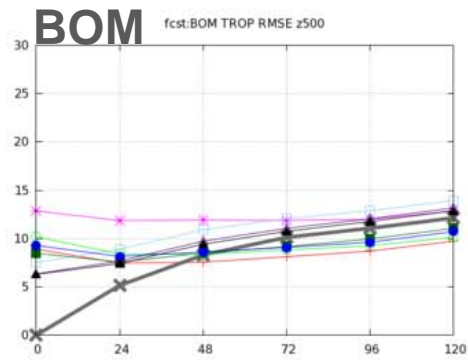


SH PSEA ACC

- Score variability is not converged even at day 5 forecast and is larger than the difference of OVN scores.
- Are there 'good' combinations such as (UKM,KMA,BOM), (JMA,CMC,NCP) and (ECM,CMC,DWD)? But not same combinations in NH.

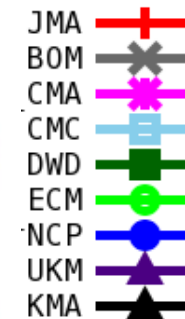
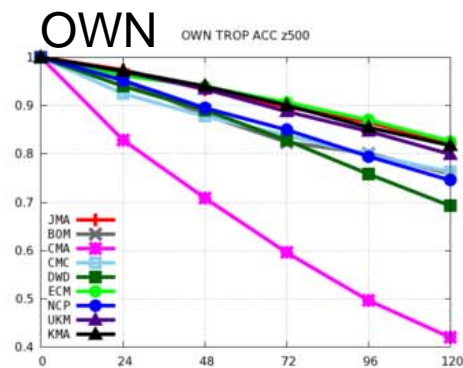
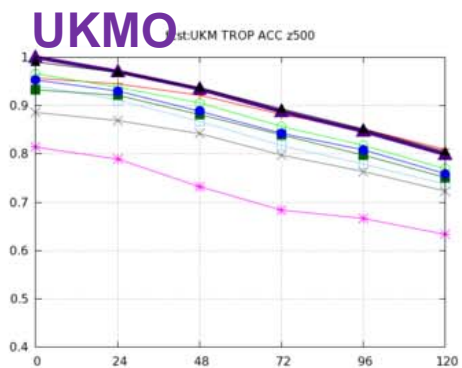
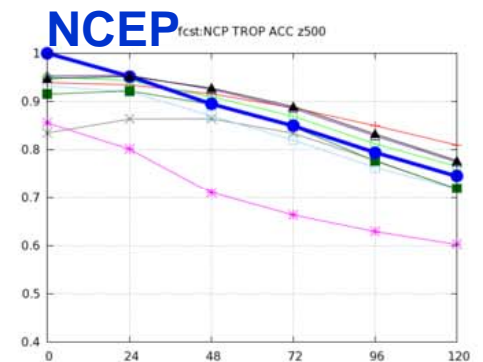
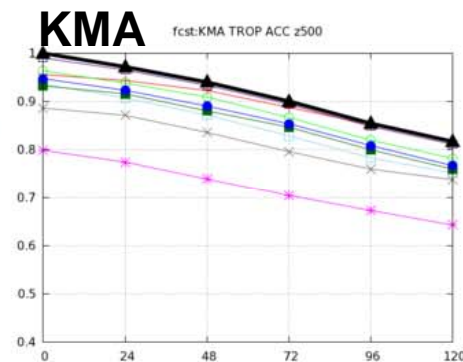
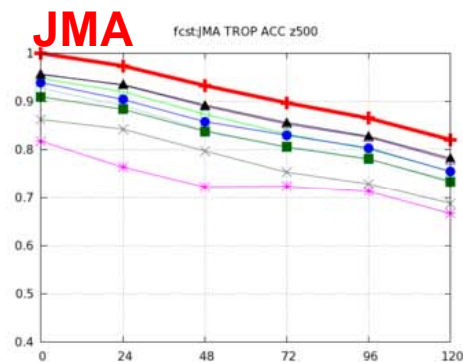
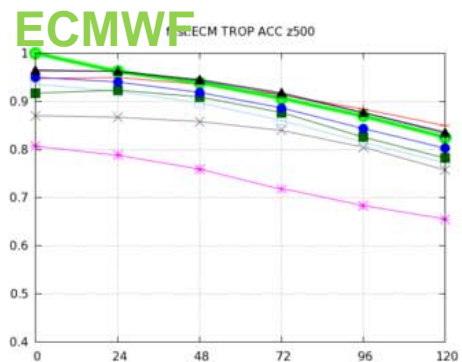
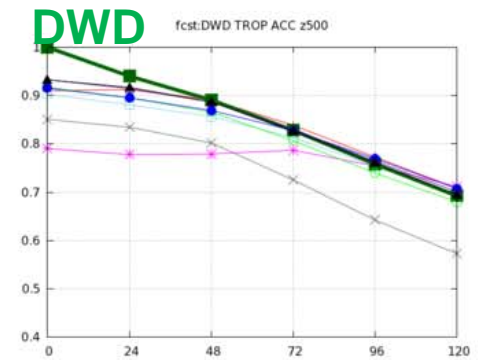
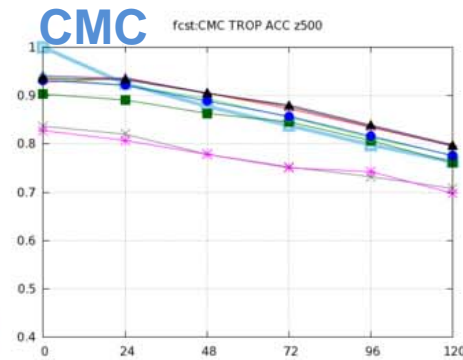
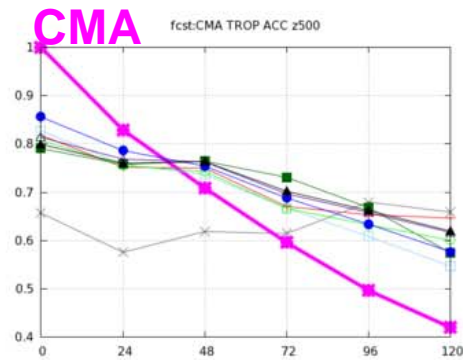
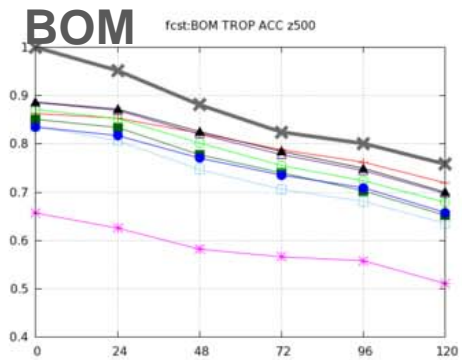
TROPICS





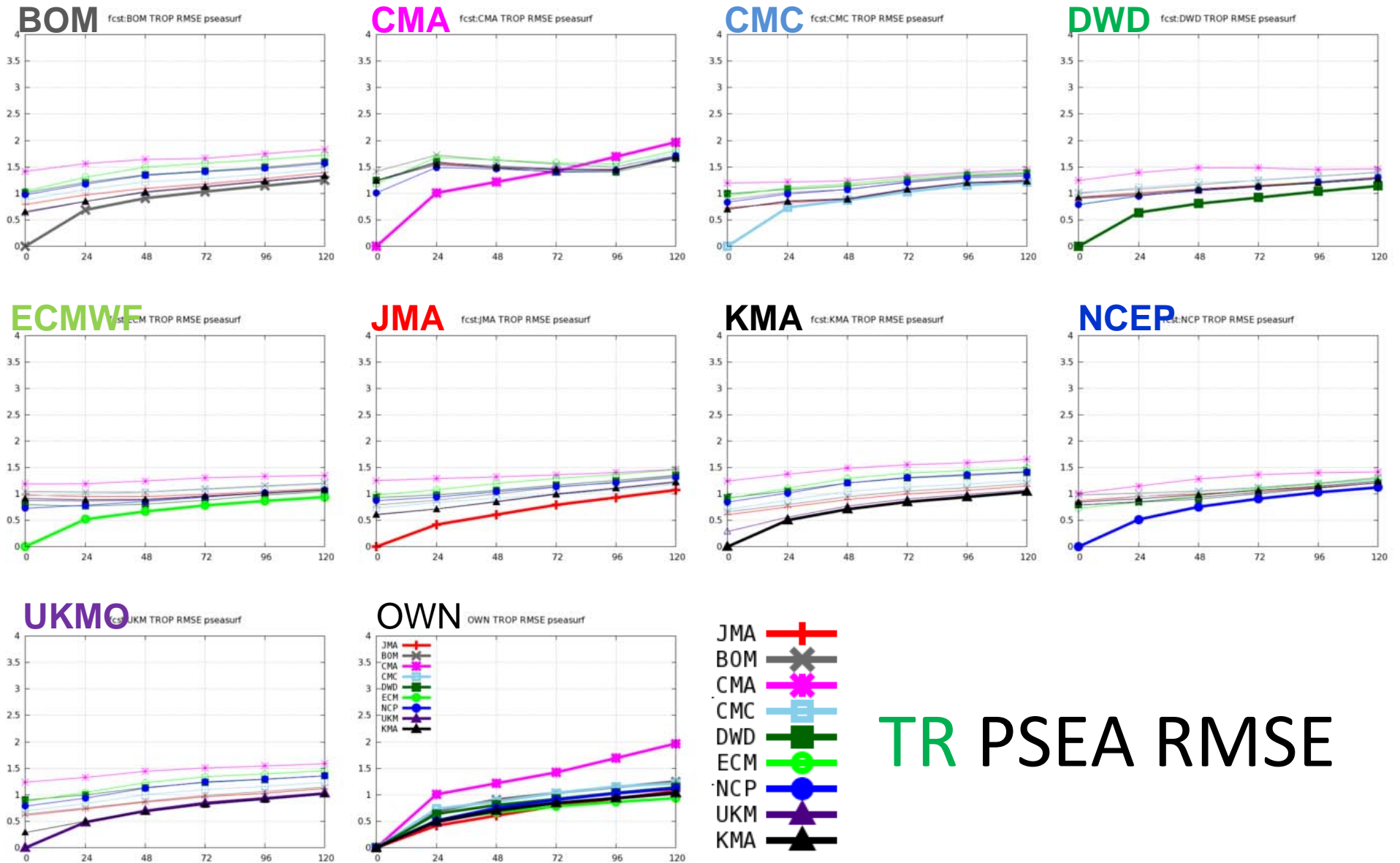
TR Z500 RMSE

- Score variability is larger than in extra-tropics.
- Score keeps unchanged, or rather decreases through forecast time and crosses own score line in some combinations.

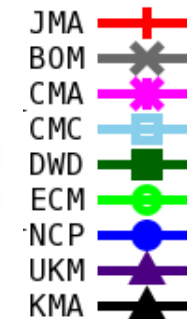
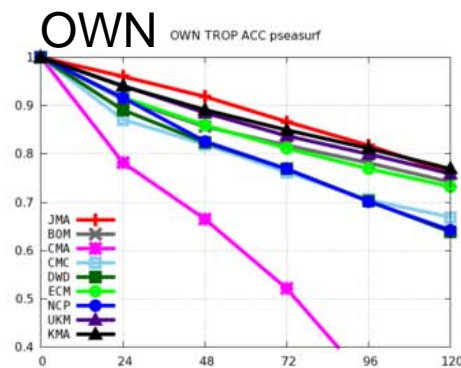
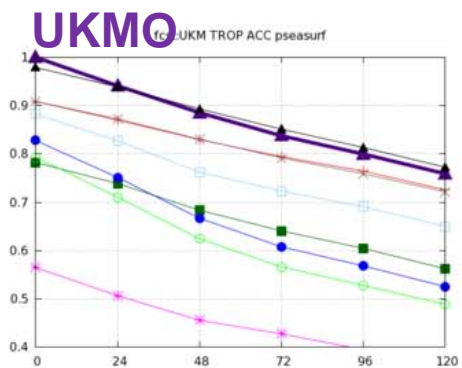
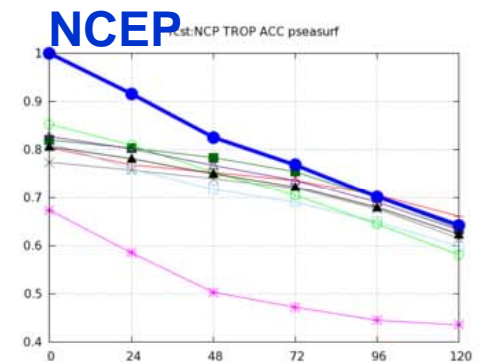
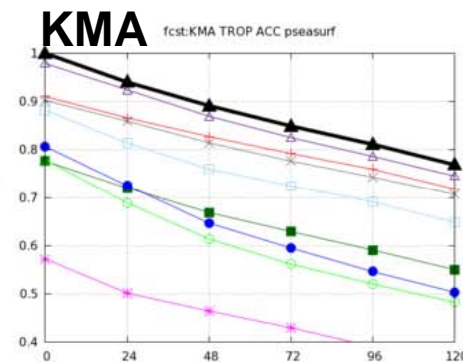
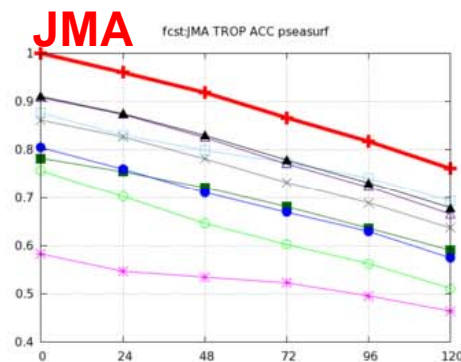
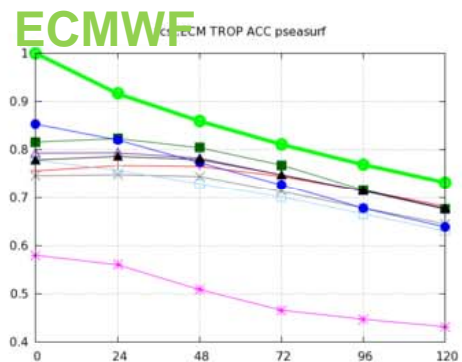
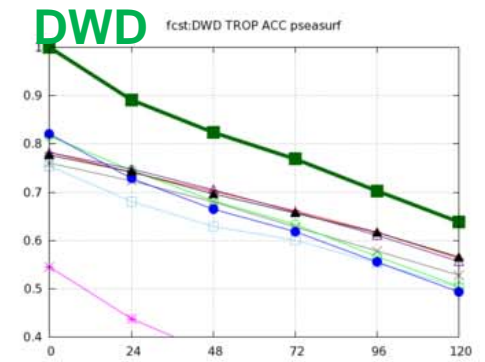
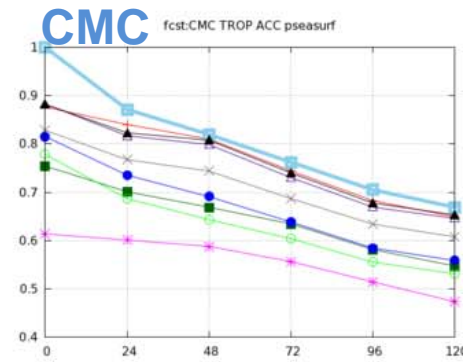
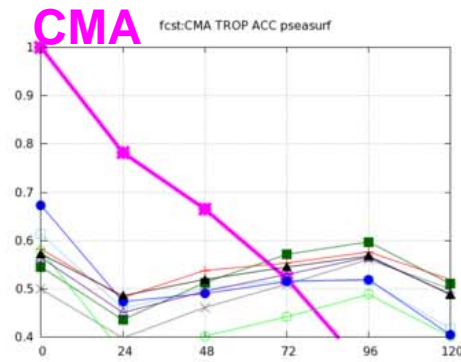
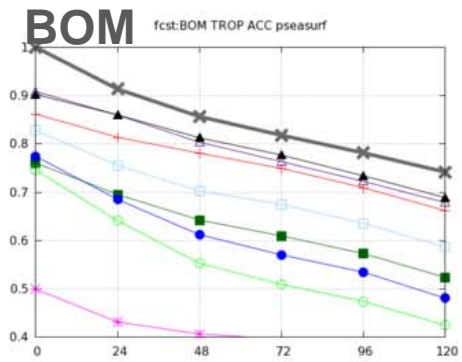


TR Z500 ACC

- Score variability is larger than in extra-tropics.



- Score variability is larger than in Z500, extra-tropics.



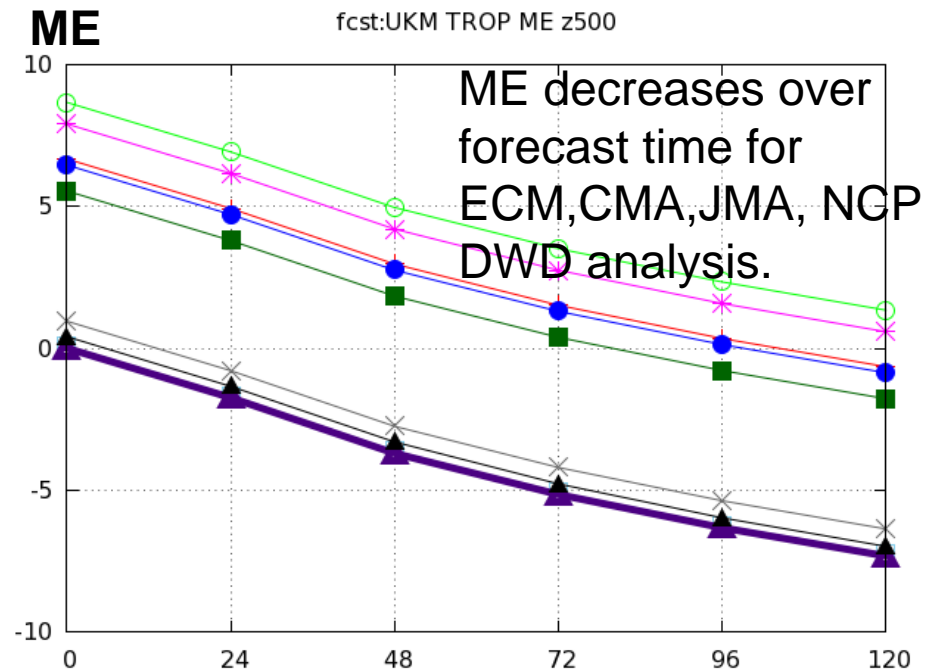
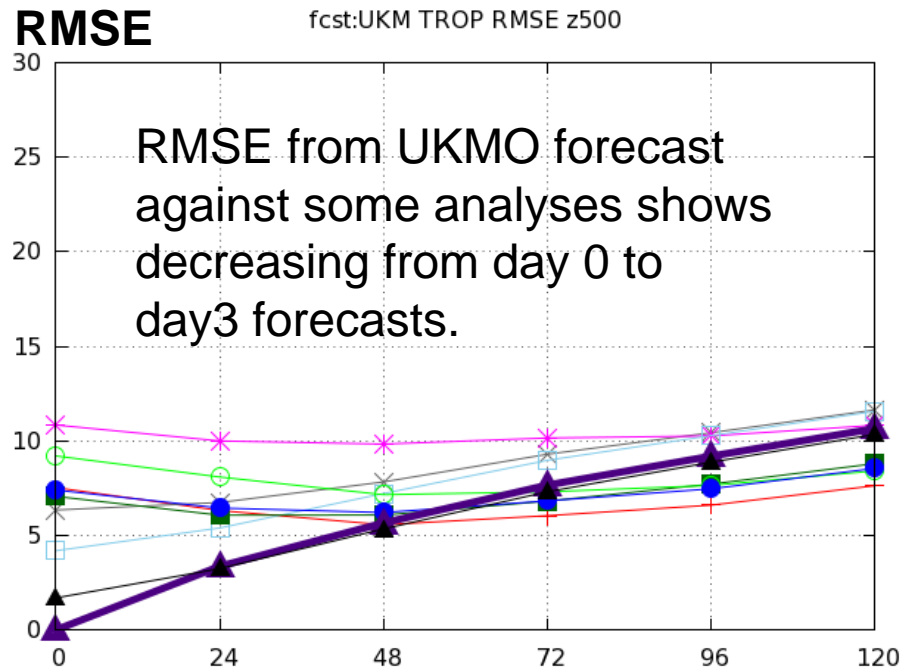
TR PSEA ACC

- Score variability is larger than in extra-tropics.
- Score keeps unchanged, or rather decreases through forecast time and crosses own score line in some combinations.



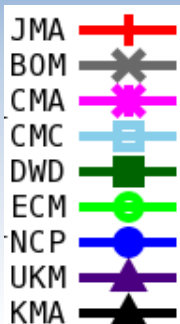
Decrease of RMSE in Tropics

Example of UKMO case.

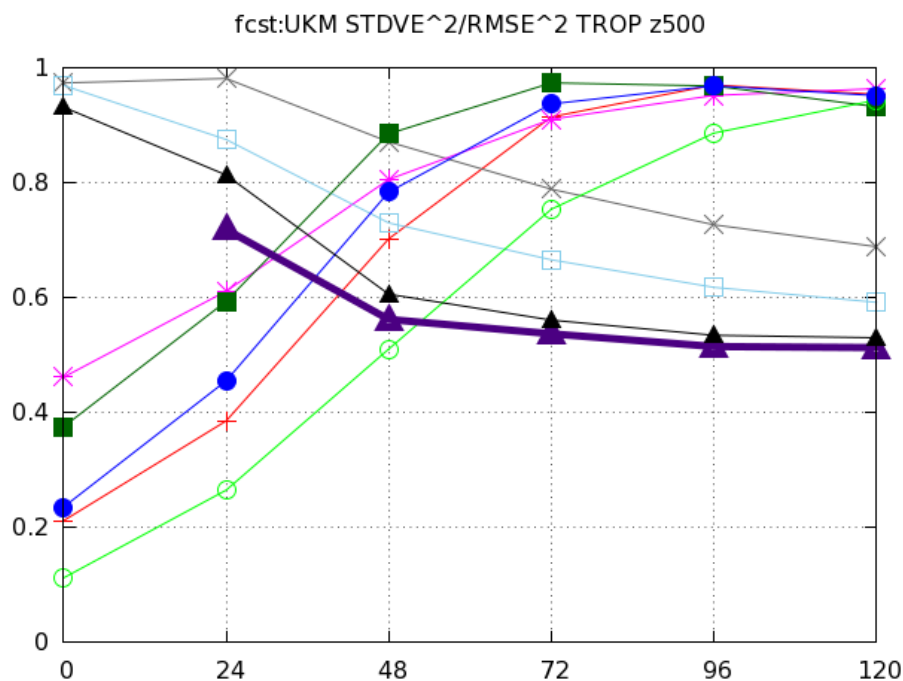


UKMO has negative forecast bias in TR.

Decrease of RMSE in Tropics



STDVE²/RMSE²



In TR, contributions of ME² to RMSE² account for 40% for RMSE against OWN analysis.

ME² contribution is becoming small for ME decreasing combinations over forecast time.


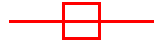
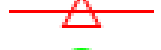


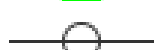



STDVE is monotonically increasing in this case (not shown).

Therefore RMSE can be decreasing over forecast time depending on forecast and analysis characteristics.

Summary of 9x9 verification

- In extra-tropics :
 - Most reliable : Z500/ACC
 - Reliable : Z500/RMSE (beyond day 2)
 - Sensitive : PSEA/ACC, RMSE and Z500/RMSE (up to day 2)
- In tropics :
 - Both of Z500 and PSEA for RMSE and ACC are sensitive and analysis-changed variability is larger than own-score variability.
 - Strange time-variation of score can be seen in some forecast-analysis combinations according to model bias.
- Among these forecast – analysis combinations, there are ‘better’ and ‘worse’ combinations which show less score appearance changes.
 - But they can differ depending on regions.

3 X 3 (JMA, ECM, BOM) T850, VW850

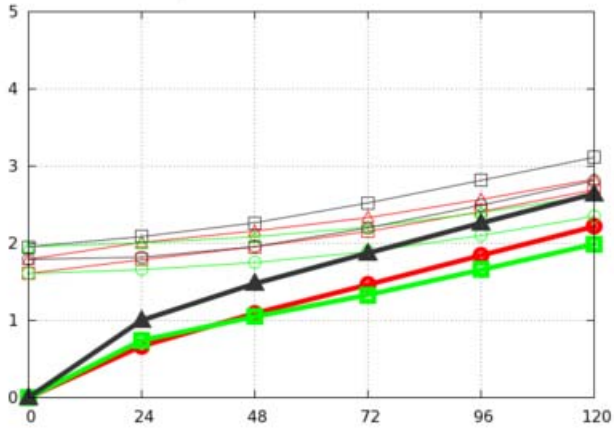
fJMA : aJMA		Forecasts are distinguished by color.
fJMA : aECM		Red : JMA
fJMA : aBOM		Green : ECMWF
fECM : aJMA		Gray : BOM
fECM : aECM		Analyses are distinguished by symbol.
fECM : aBOM		Circle : JMA
fBOM : aJMA		Square : ECM
fBOM : aECM		Triangle : BOM
fBOM : aBOM		

- Verification against own analysis is drawn with a **THICK** line.
- Score
 - T850 : RMSE, ACC
 - VW850 : RMSVE, ACC(WS)
- Y-range of RMSE(RMSVE) in TR is half of those in NH and SH.

T850

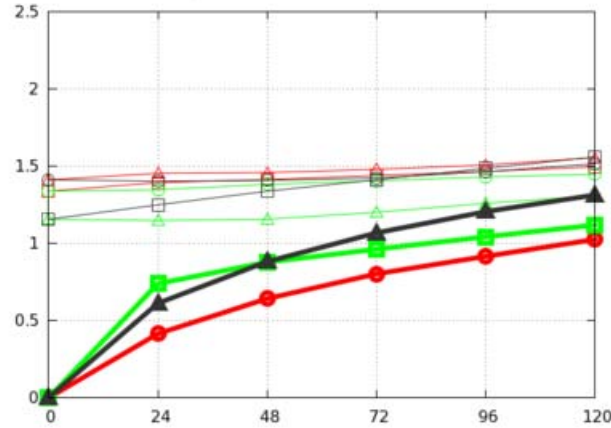
NH

JMA-ECM-BOM NHEM RMSE t850



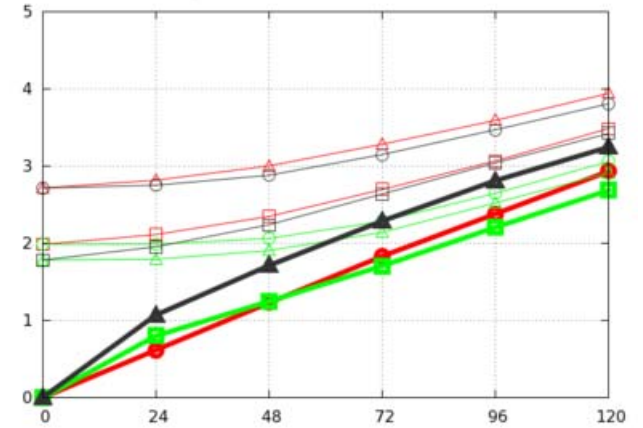
TR

JMA-ECM-BOM TROP RMSE t850



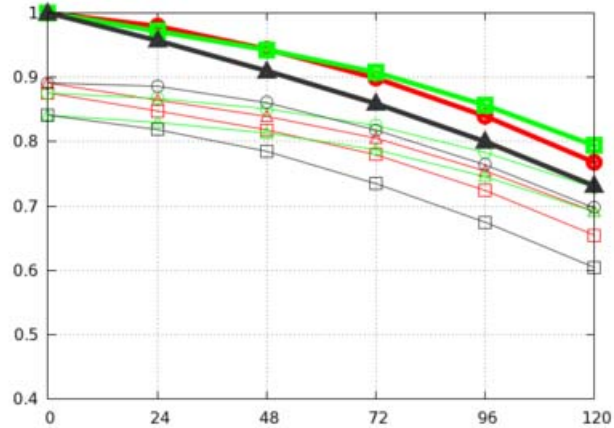
SH

JMA-ECM-BOM SHEM RMSE t850

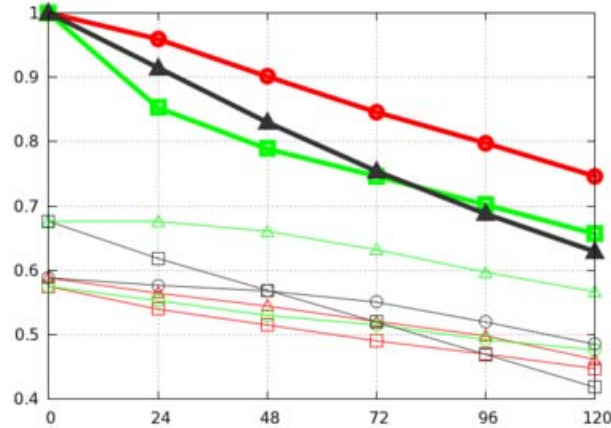


RMSE

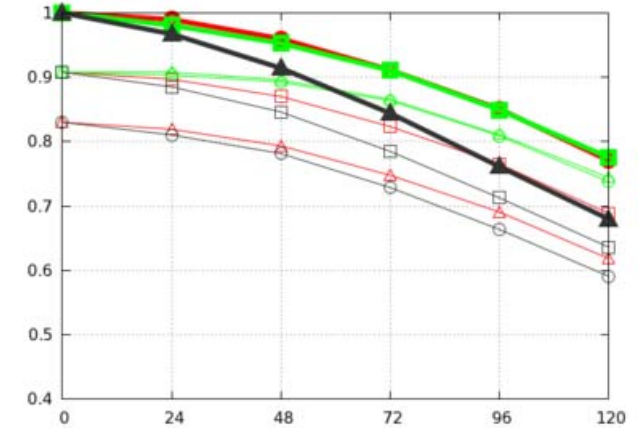
JMA-ECM-BOM NHEM ACC t850



JMA-ECM-BOM TROP ACC t850

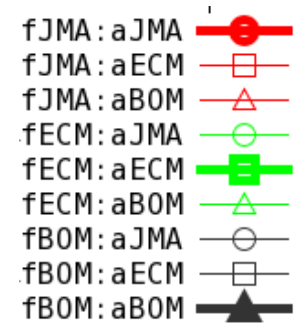


JMA-ECM-BOM SHEM ACC t850



ACC

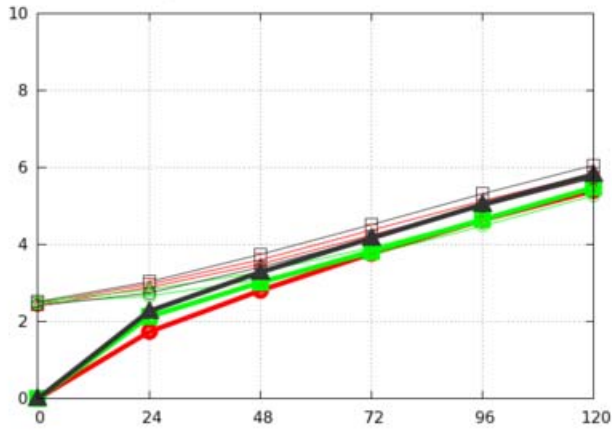
- Differences between analysis-changed scores and OWN-analysis scores are not converged even at day 5 in all regions .



VW850

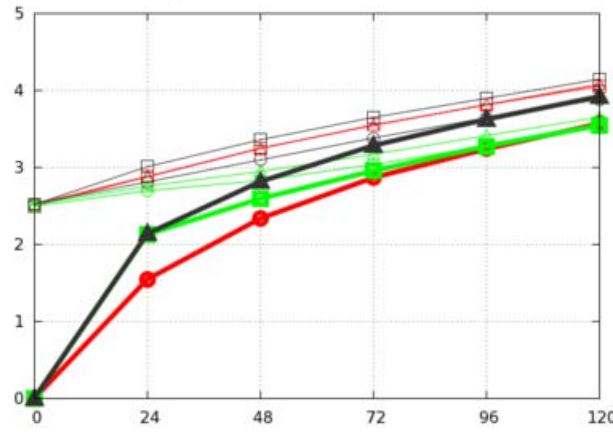
NH

JMA-ECM-BOM NHEM RMSVE vw850



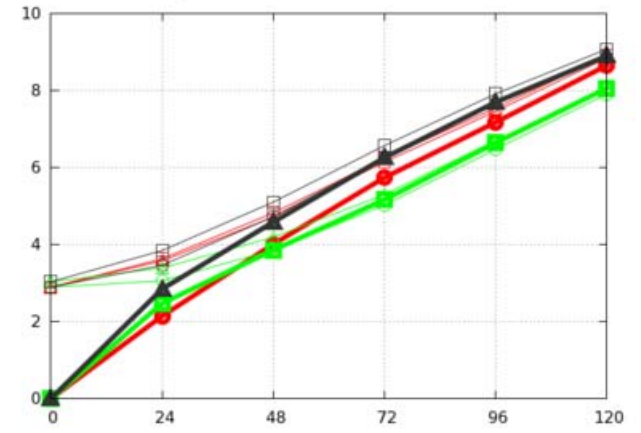
TR

JMA-ECM-BOM TROP RMSVE vw850



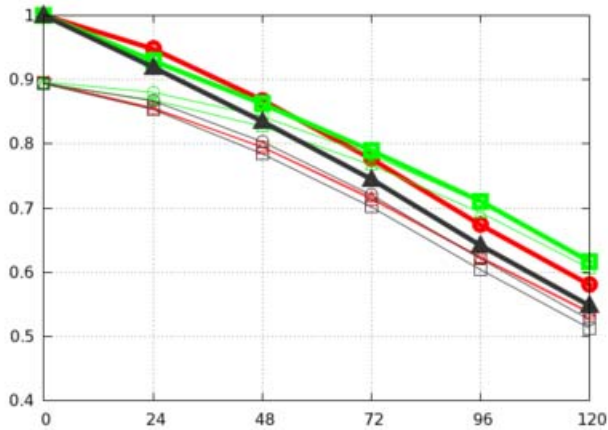
SH

JMA-ECM-BOM SHEM RMSVE vw850

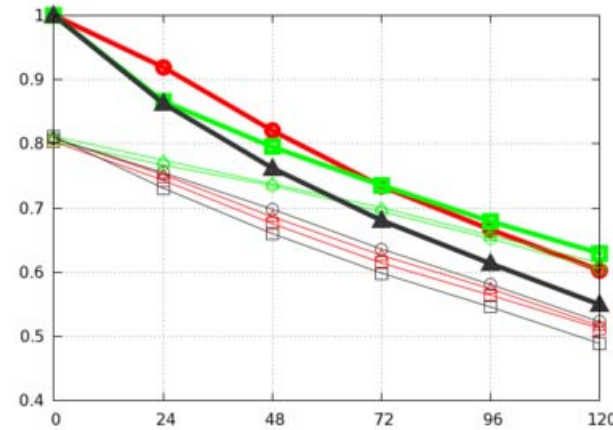


RMSVE

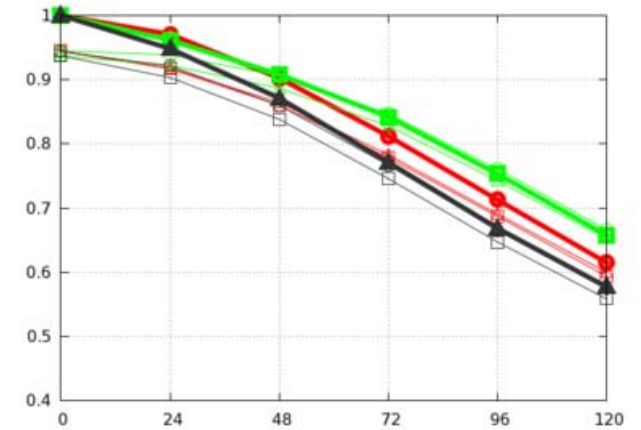
JMA-ECM-BOM NHEM ACC vw850



JMA-ECM-BOM TROP ACC vw850

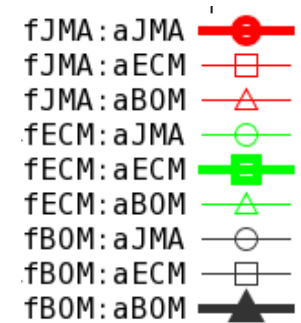


JMA-ECM-BOM SHEM ACC vw850



ACC

- ECMWF forecasts show good convergence even if verified against other analyses.
- In NH and SH, differences of RMSVE are converged at day 5, but ACC are not.
- In TR, analysis-changed scores look different from own-analysis scores.



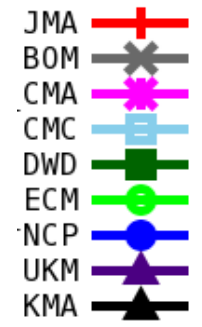
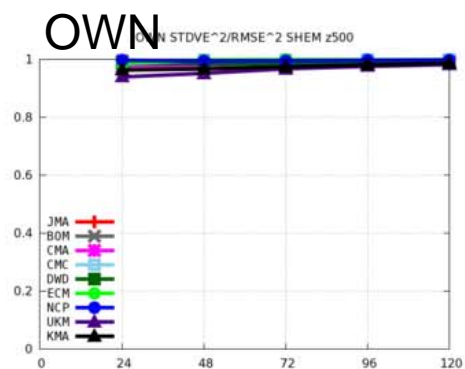
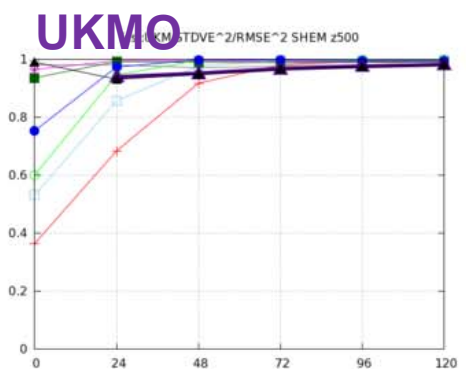
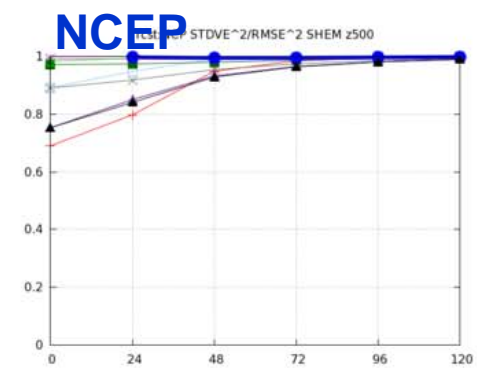
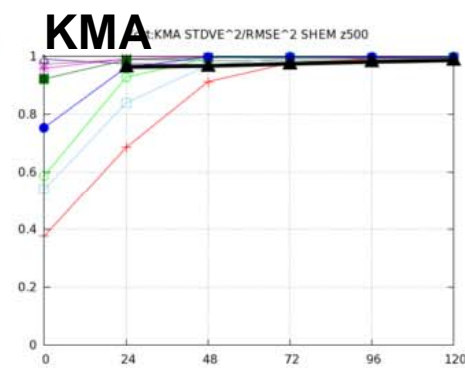
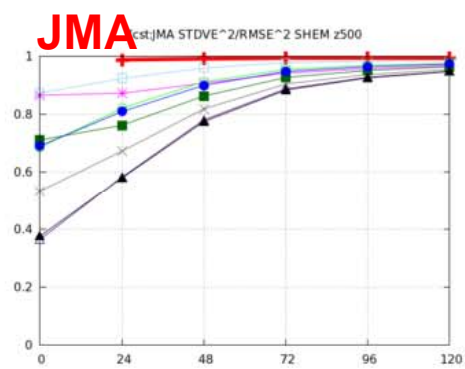
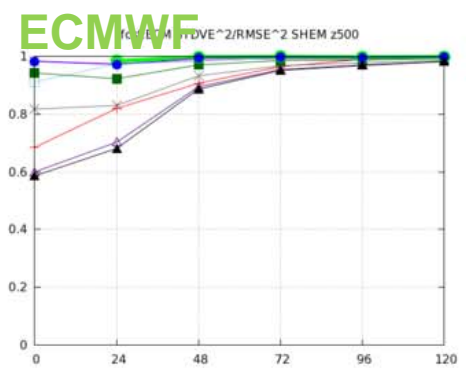
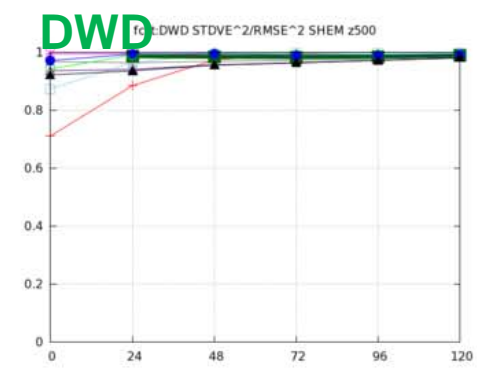
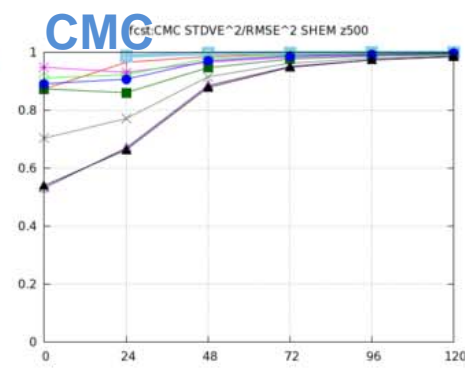
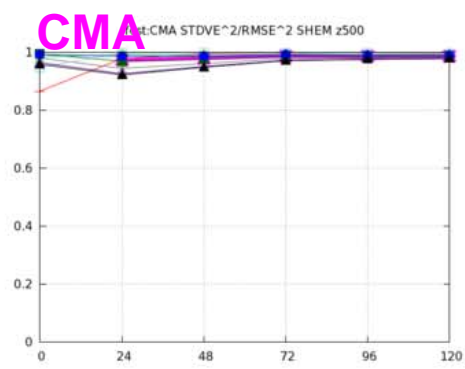
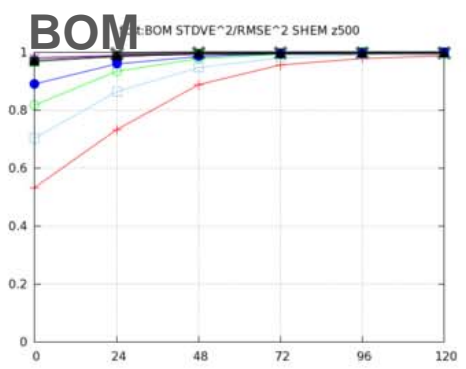
Total Summary

- In extra-tropics :
 - Z500/ACC is most reliable score to measure forecast accuracy.
 - For over day 2 forecast ranges, Z500/RMSE and VW850/RMSVE, ACC can be used.
 - But careful comparison and interpretations are necessary.
 - Score of PSEA, T850 are sensitive to choice of analysis.
- In tropics :
 - score behavior looks strange. The influences of analysis fields to verification results are large and differences between analyses are not negligible when we compare scores verified against own analysis.

BACKUP

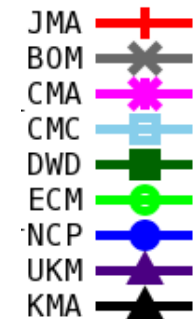
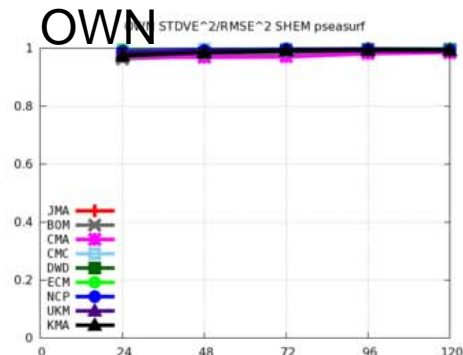
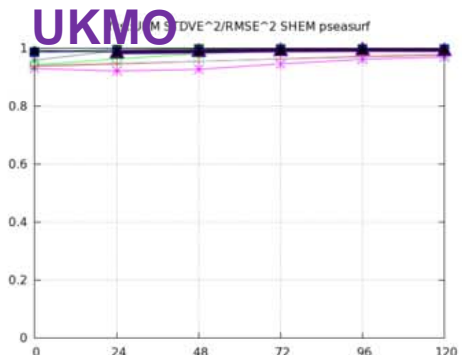
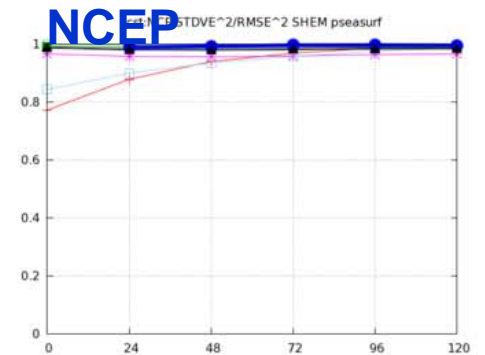
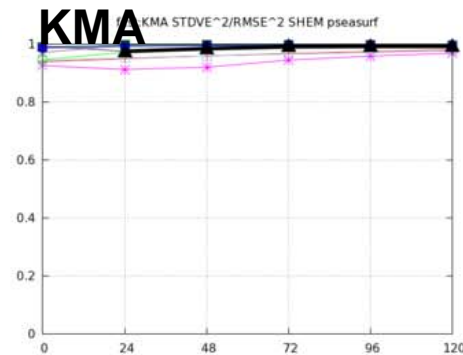
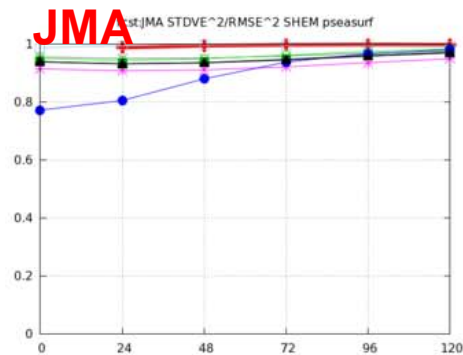
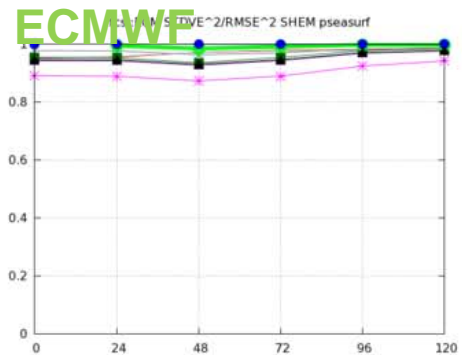
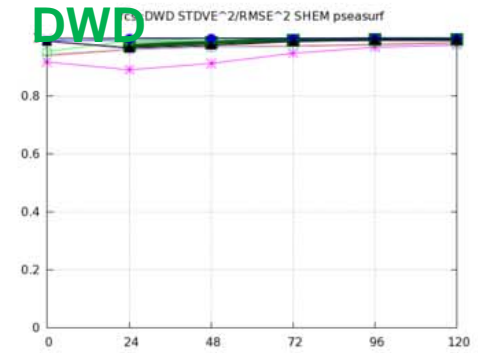
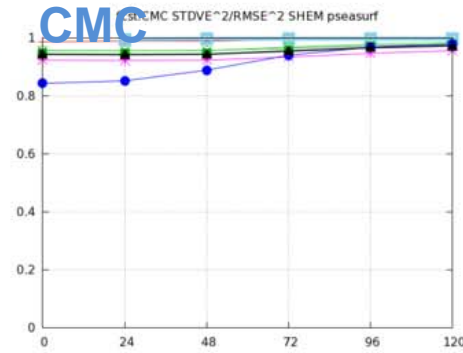
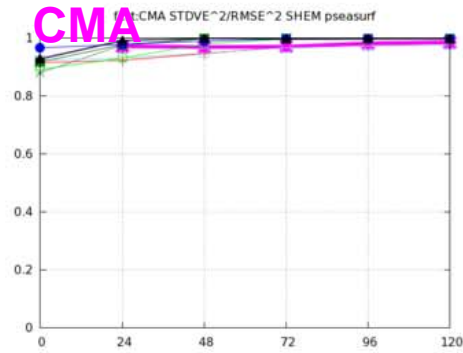
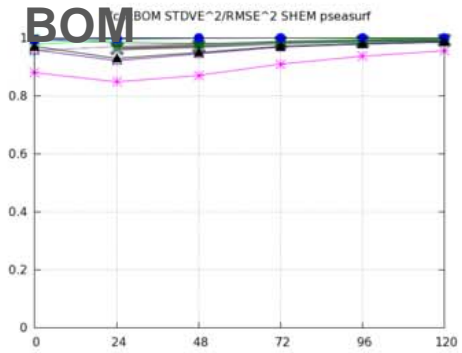
SOUTHERN HEMISPHERE





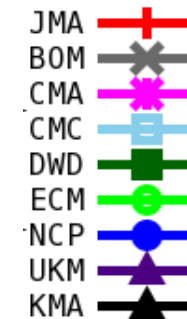
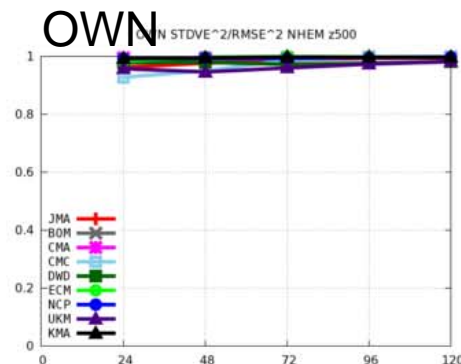
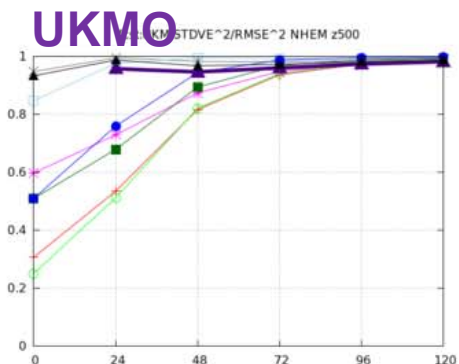
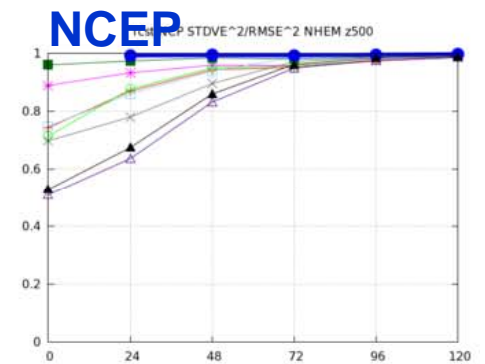
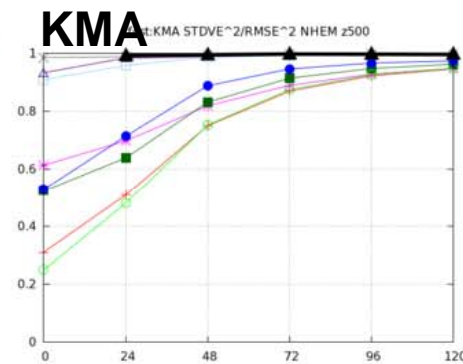
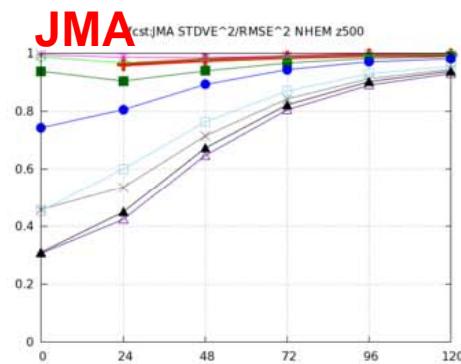
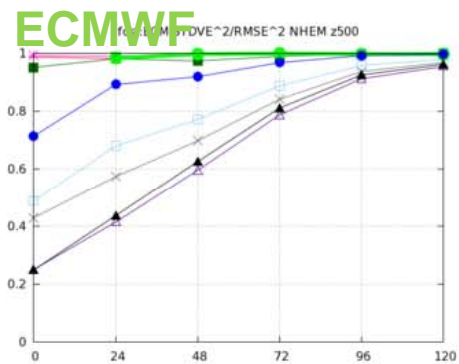
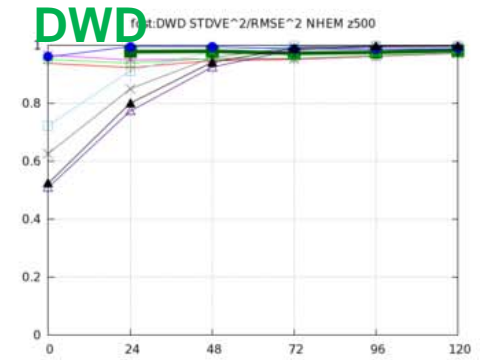
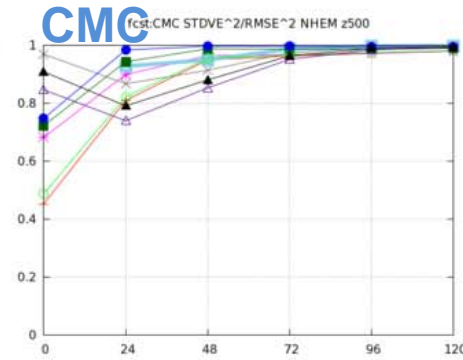
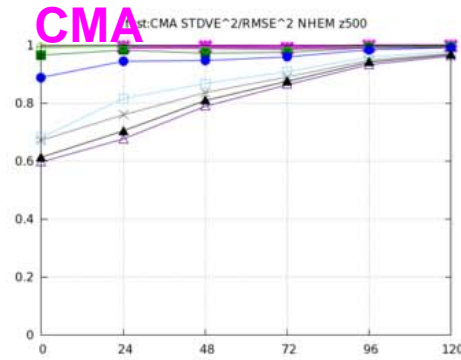
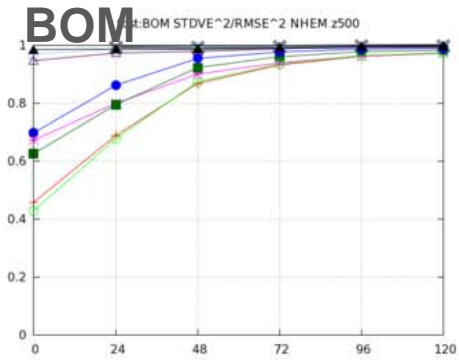
SH Z500
 STDVE²/RMSE²

• aaaa

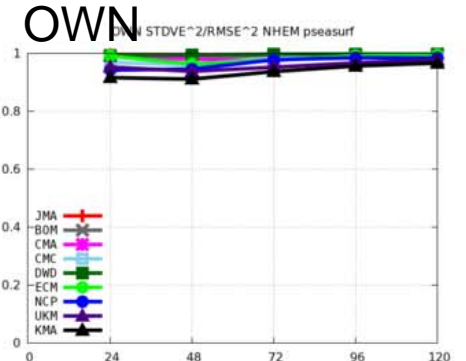
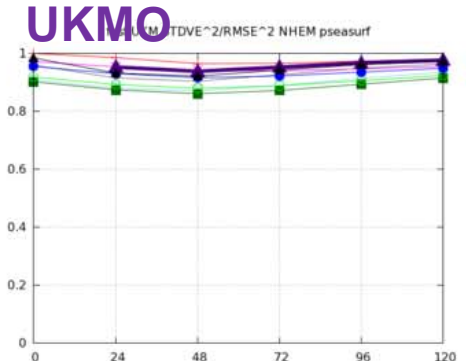
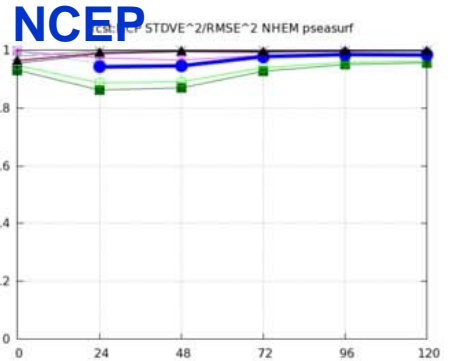
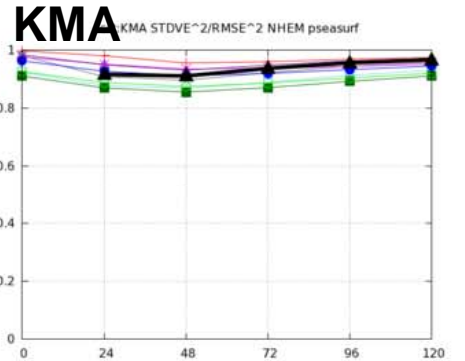
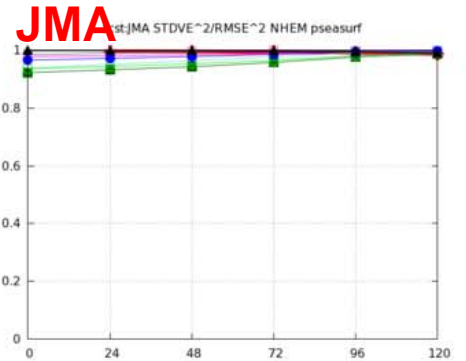
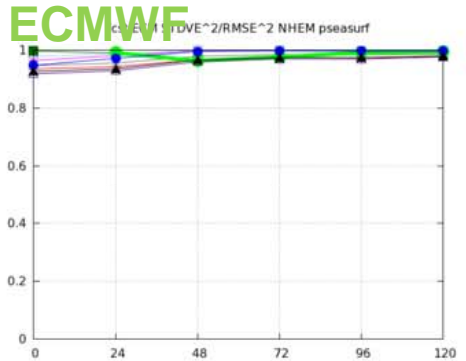
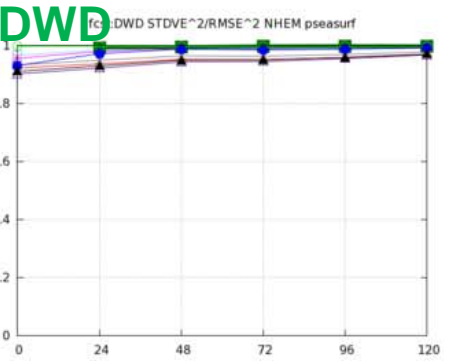
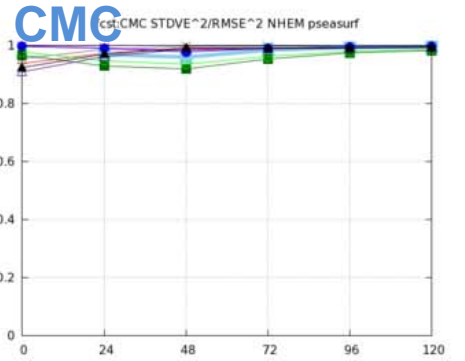
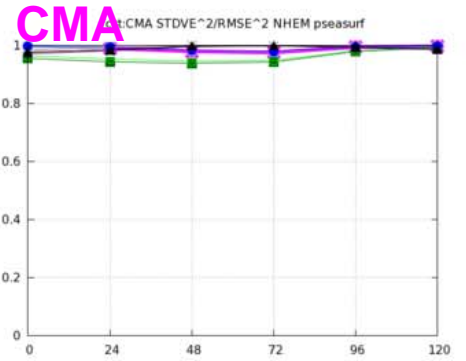
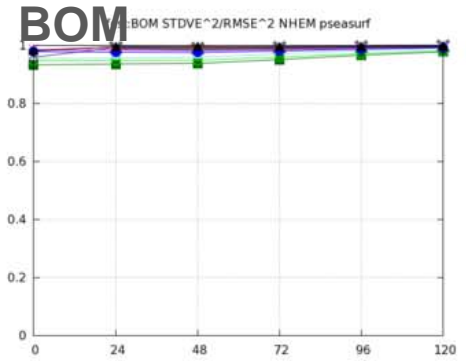


SH PSEA
 STDVE²/RMSE²

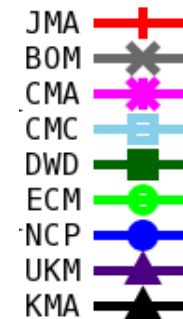
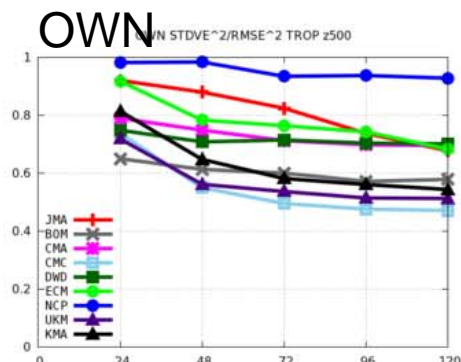
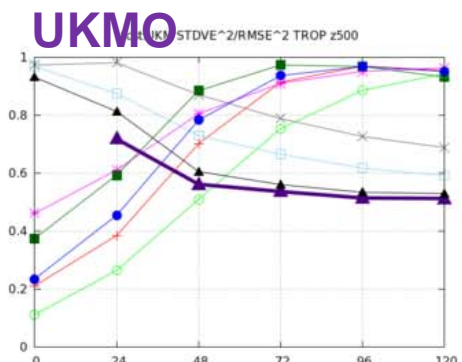
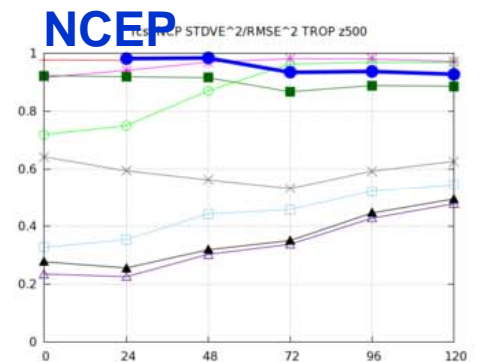
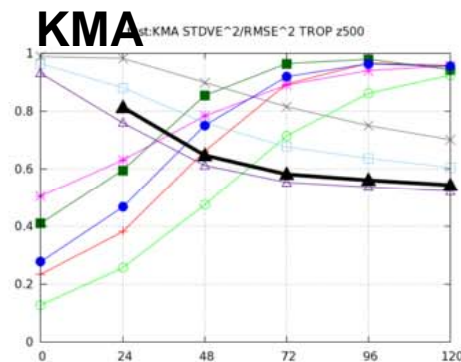
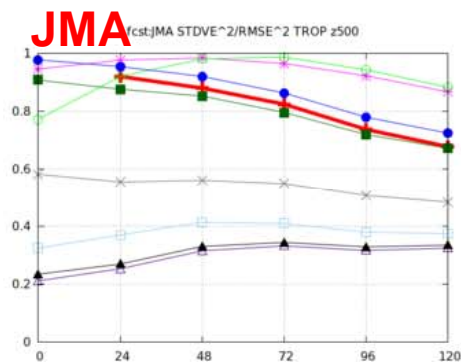
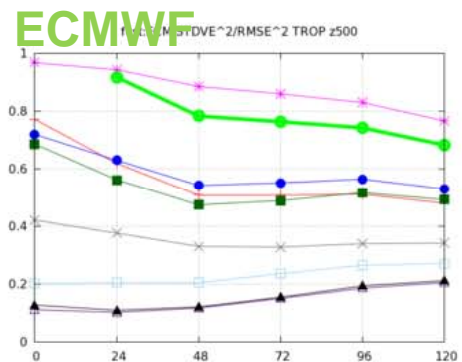
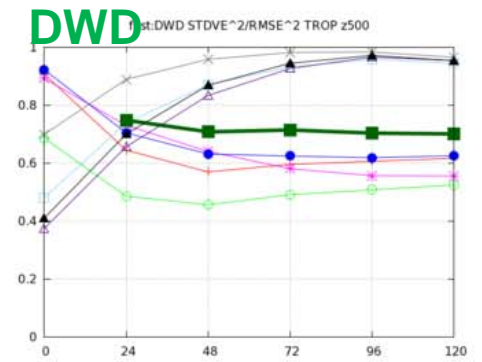
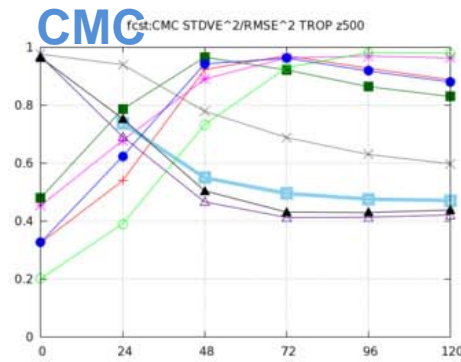
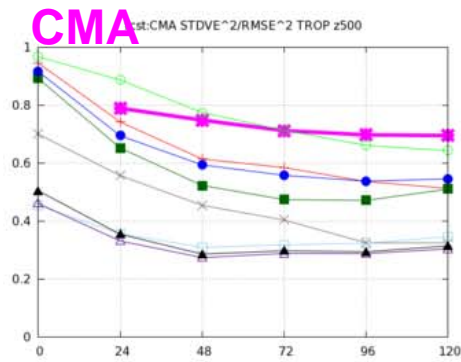
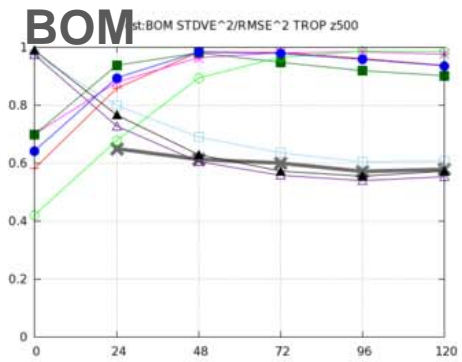
• aaaa



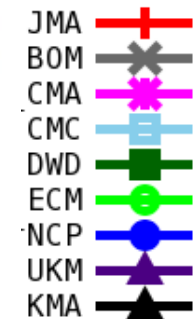
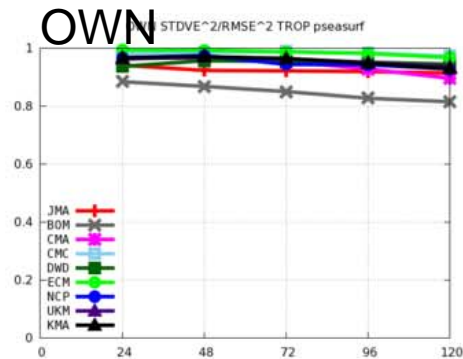
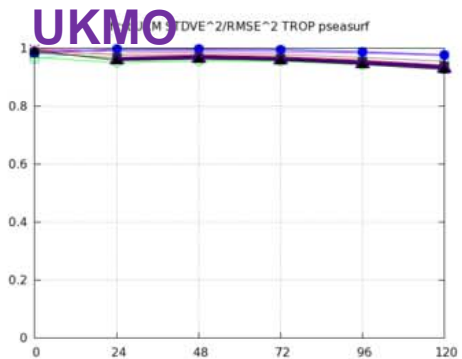
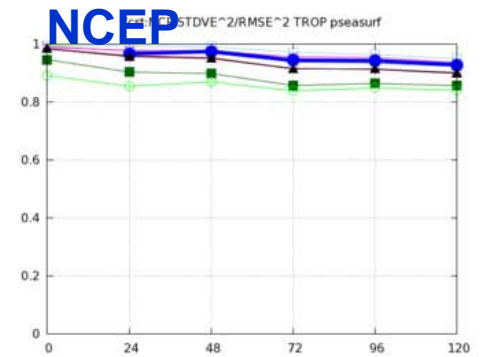
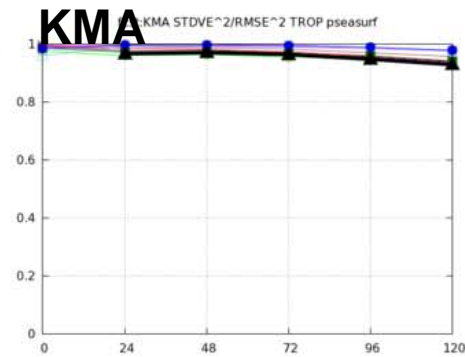
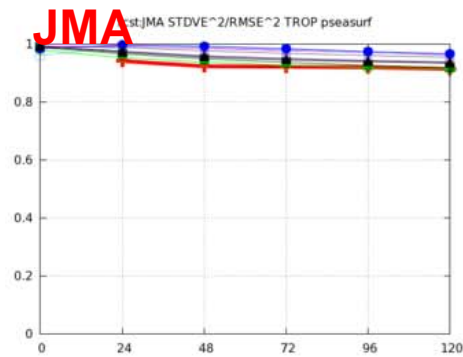
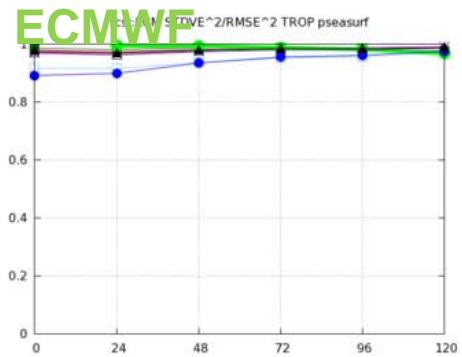
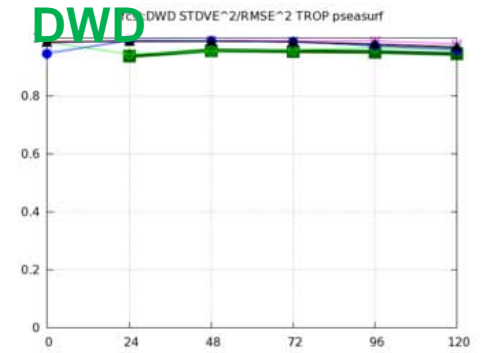
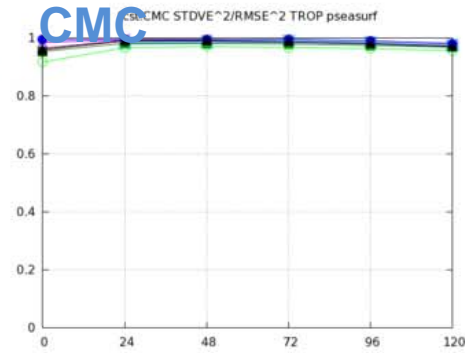
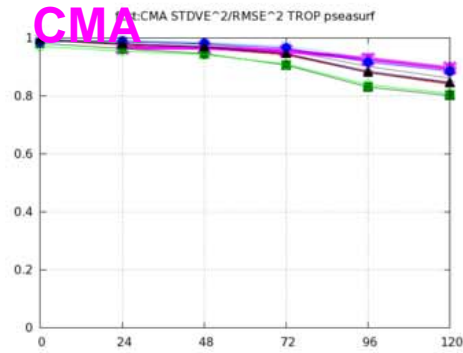
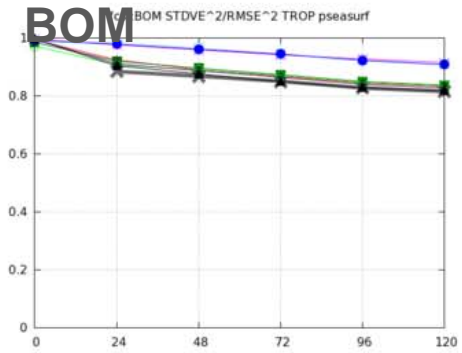
NH Z500
STDVE²/RMSE²



NH PSEA
 STDVE²/RMSE²



TR Z500
STDVE²/RMSE²



TR PSEA
 STDVE²/RMSE²