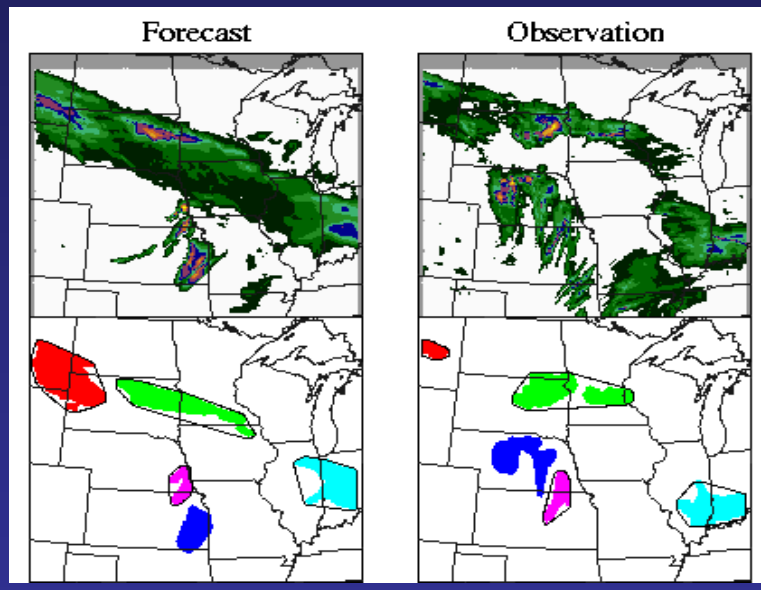
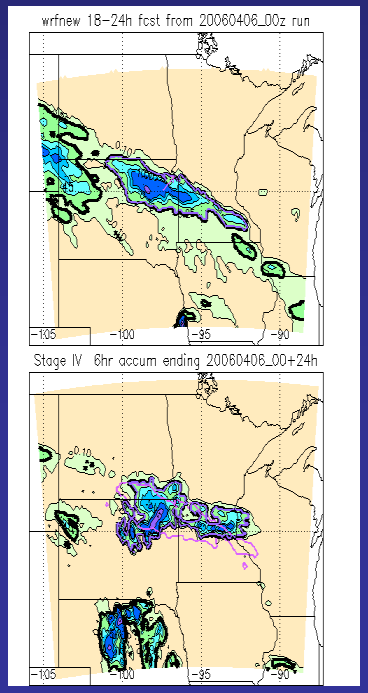
Object-oriented verification: Application to ensemble precipitation forecasts

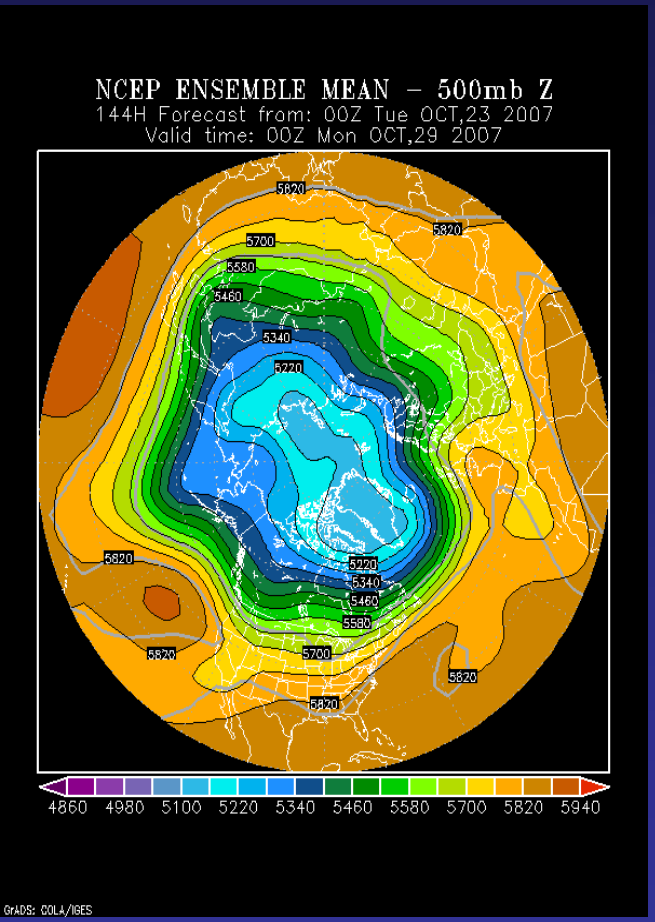


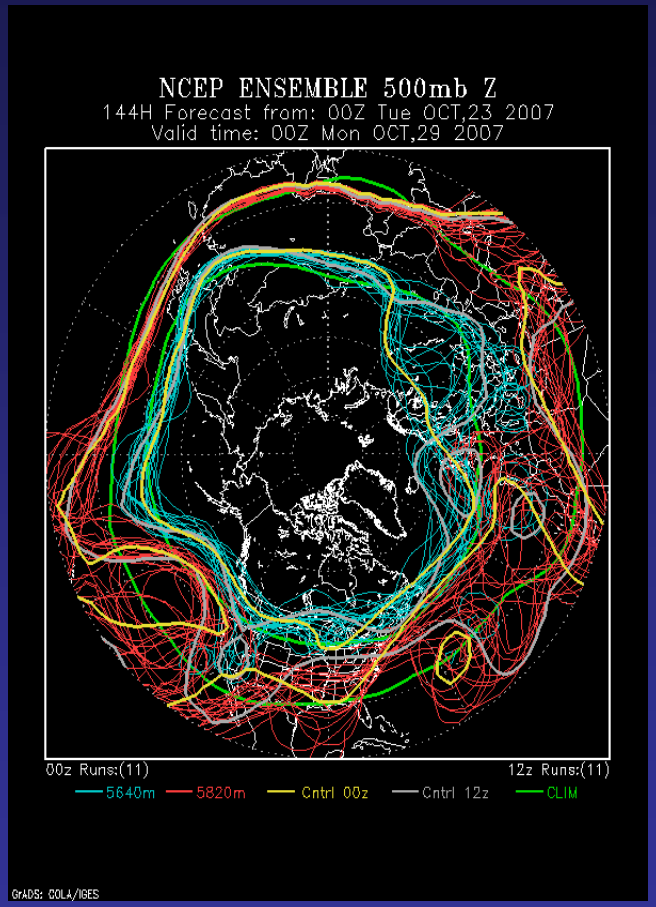


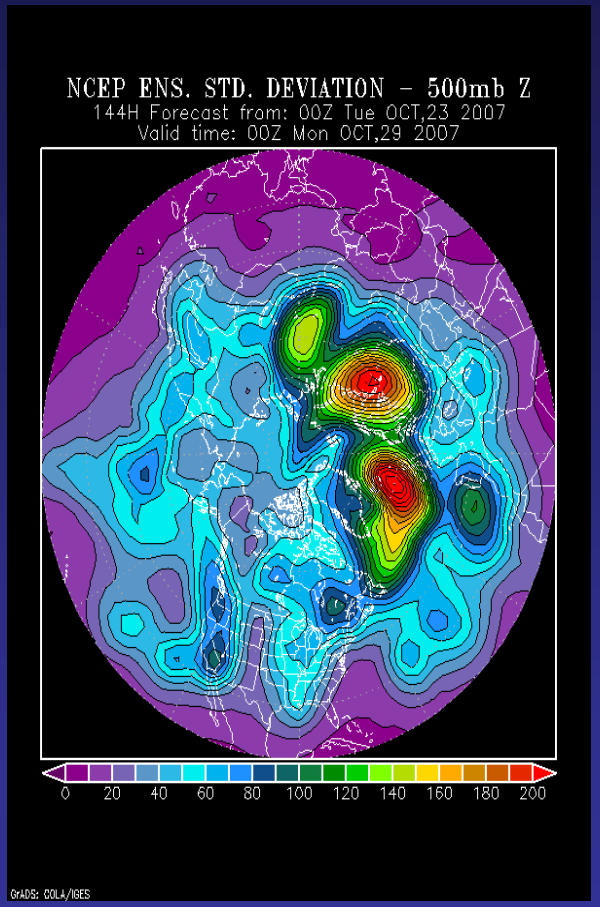
William A. Gallus, Jr.

Iowa State University

Ensemble Forecasts - the primary NWP tool of the future?





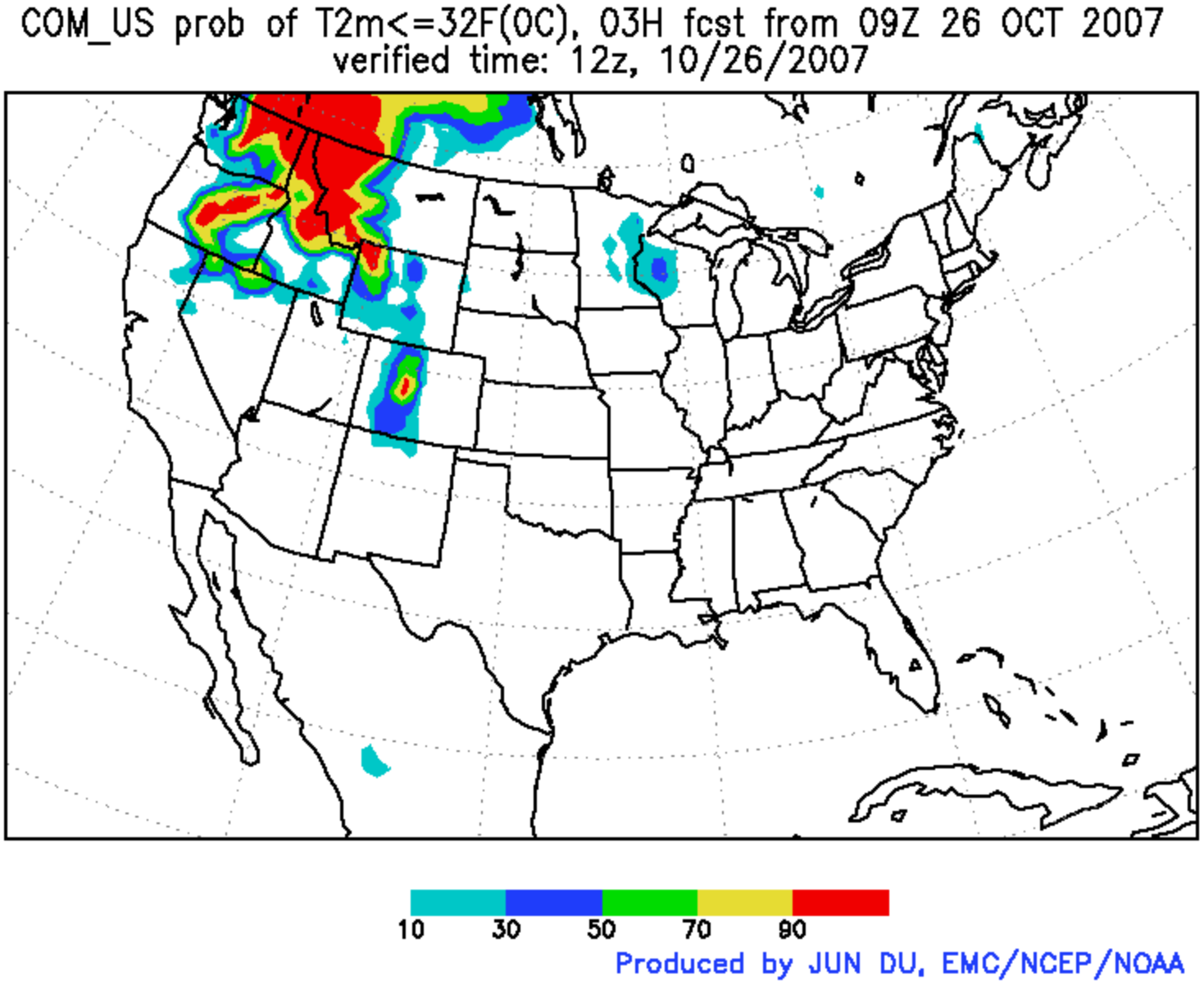


Mean

Spaghetti Plots

Standard Deviation (uncertainty)

Or…. Probability information



How can object-oriented techniques be applied to ensemble forecasts?

- Evaluation of ensemble mean forecasts (really just another deterministic forecast)
- Evaluation of spread of object parameters – does this relate to skill?
- Does the mean of the various object parameters from each member give a more accurate depiction of the event than any specific deterministic forecast?
- Might object-oriented approaches be applied to probability forecasts?

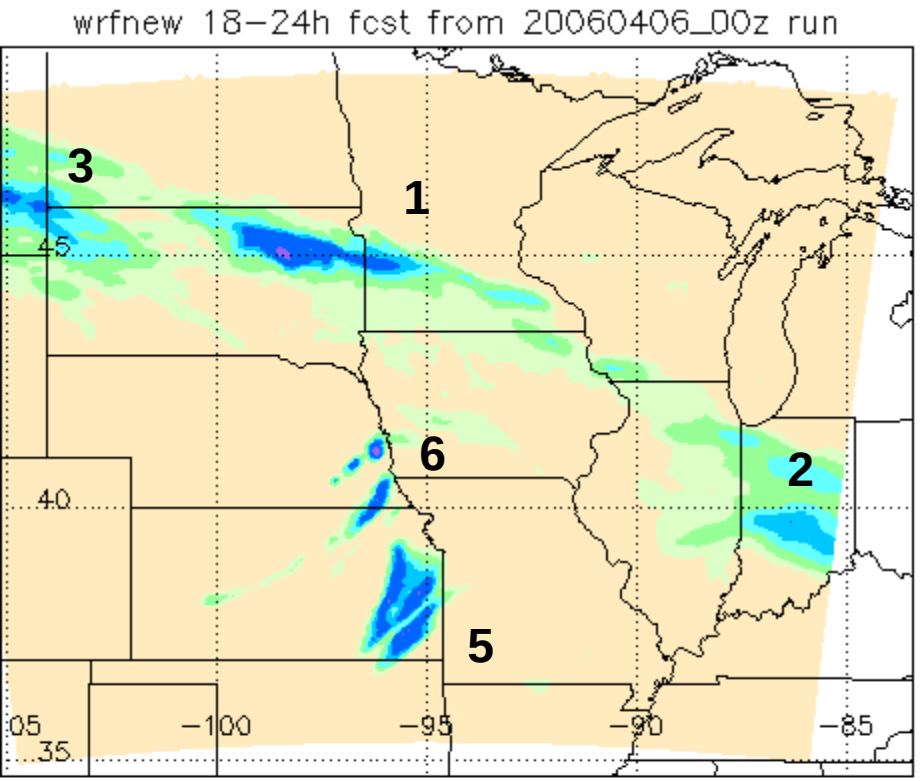
Object-oriented technique parameters:

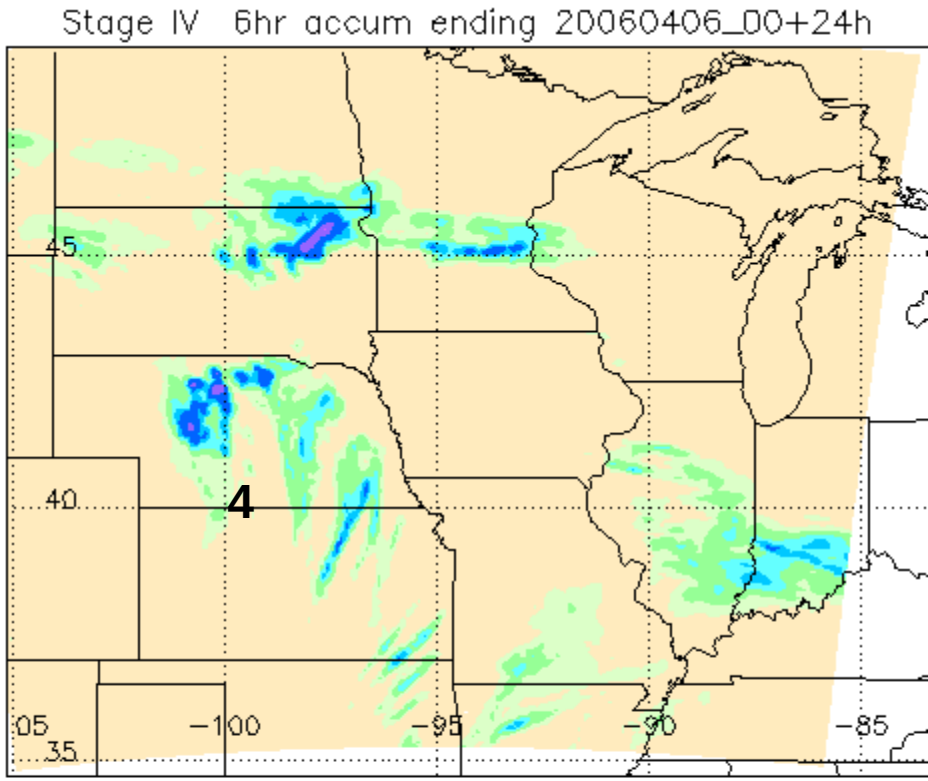
- Areal coverage
- Mean Rain Rate
- Rain Volume of system
- Number of systems
- Displacement error
- Axis orientation
- Correspondence Ratio
- Correlation coefficients/RMSE
- Etc….

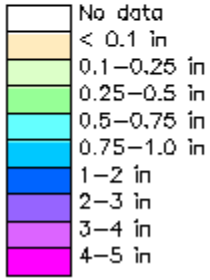
Methodology

- 6h 15-km grid precipitation forecasts over 60 h integration periods from TWO 8-member WRF ensembles run by Adam Clark for 72 cases were input into CRA and MODE
- TWO ensembles – one with just mixed physics, the other with just perturbed I.C. and L.B.C
- This results in 10 x 16 x 72 = 11,520 evaluations (plots, tables….) from each approach

Example of 6 hr precip forecast (left) and obs (right) with CRA-derived systems indicated







	Forecast <0.25	Forecast ≥0.25
<0.25	21124	1891
≥0.25	1635	531

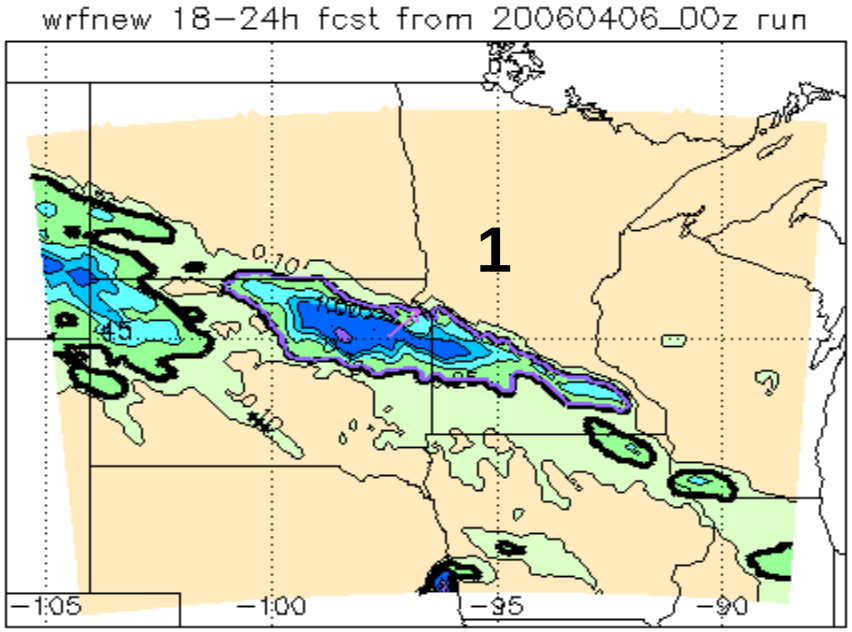
Validation statistics for 20060407_24Z Valid gridpoints=25181 Threshold=0.25 in

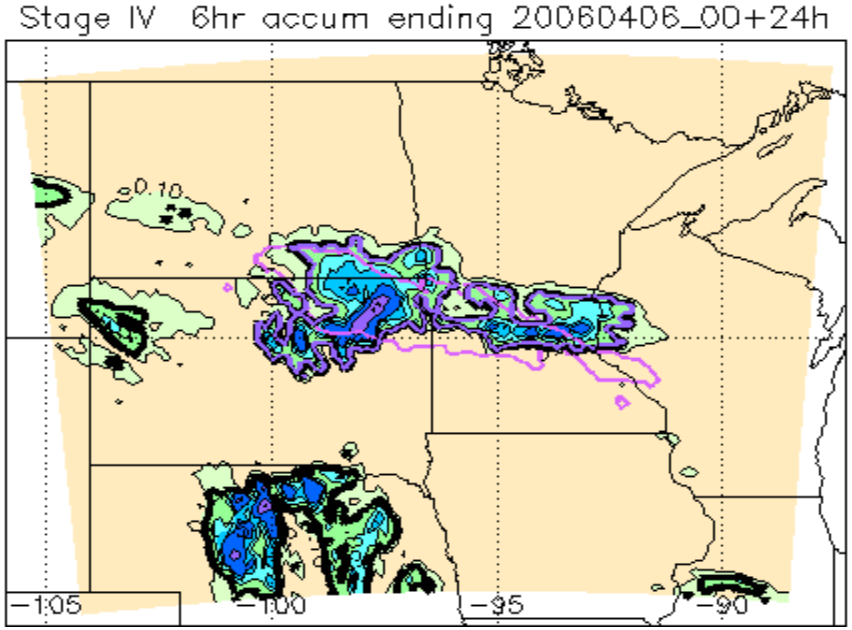
	Analysed	Forecast
# gridpoints raining	2166	2422
Average rainrate (in/ 6 hr)	0.0714	0.0801
Rain volume (km^3)	4.45	4.99
Maximum rain (in/ 6 hr)	3.19	3.46

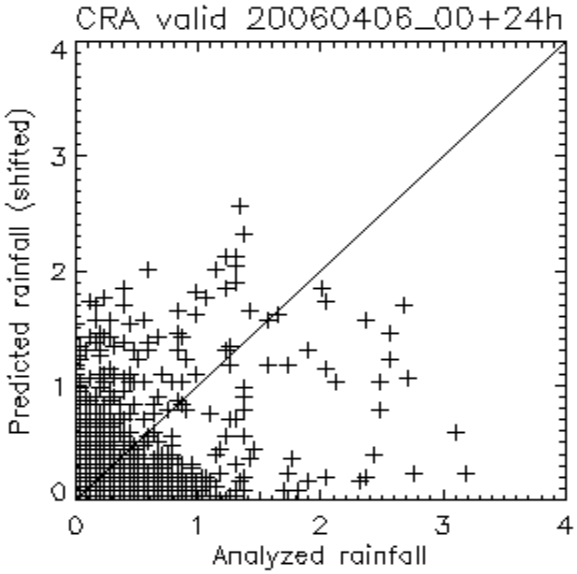
Mean abs error = 0.1088 in/ 6 hr
RMS error = 0.2526 in/ 6 hr
Correlation coeff = 0.218
Bias score = 1.118
Probability of detection = 0.245
False alarm ratio = 0.781
Hanssen & Kuipers score = 0.163
Equitable threat score= 0.084

Ebert and McBride, 2000

CRA Output for one "CRA" from Apr 6, 2007 18-24 hr forecast







wrfnew 18−24h fcst from 20060406_00z run
(43.78°,−101.05°) to (46.60°,−91.98°)
n=1089 CRA threshold=0.25 in/ 6 hr

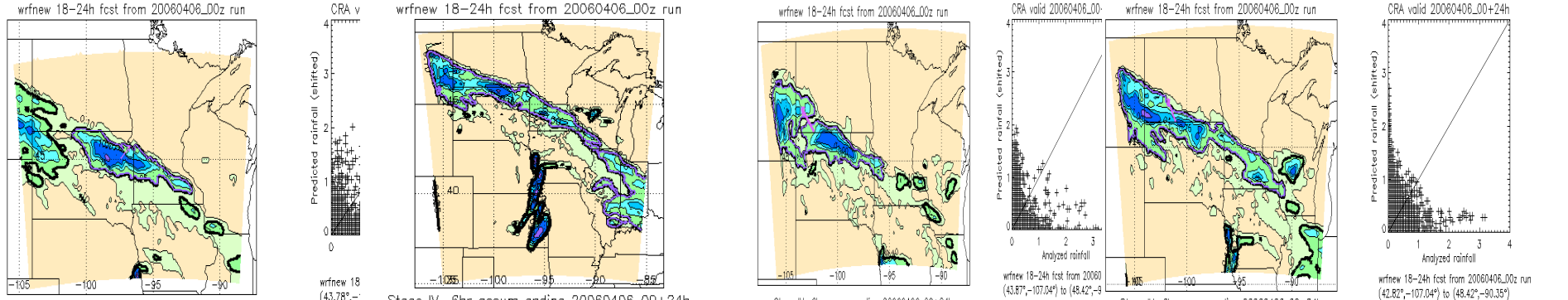
	Analysed	Forecast
# gridpoints ≧0.25 in/ 6 hr	583	534
Average rainrate (in/ 6 hr)	0.67	0.69
Maximum rain (in/ 6 hr)	3.19	2.56
Rain volume (km^3)	0.92	0.87

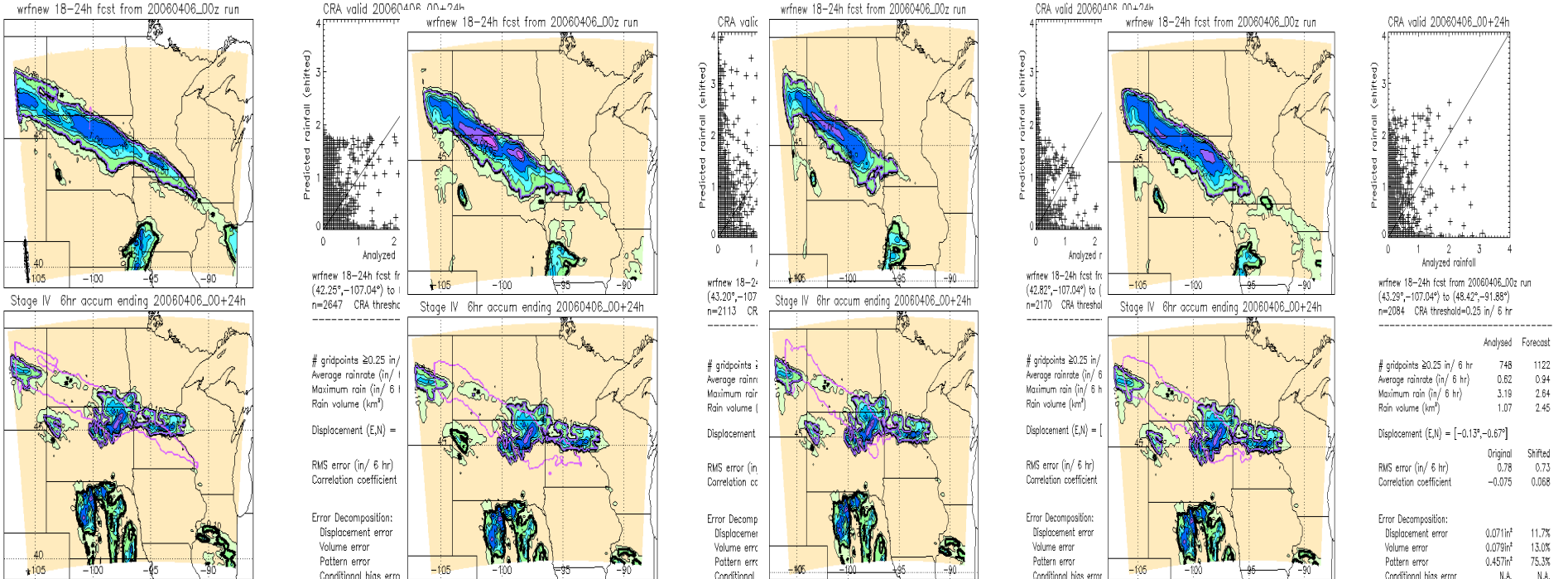
Displacement (E,N) = [−0.66°,−0.38°]

	Original	Shifted
RMS error (in/ 6 hr)	0.61	0.48
Correlation coefficient	0.063	0.390

Error Decomposition:

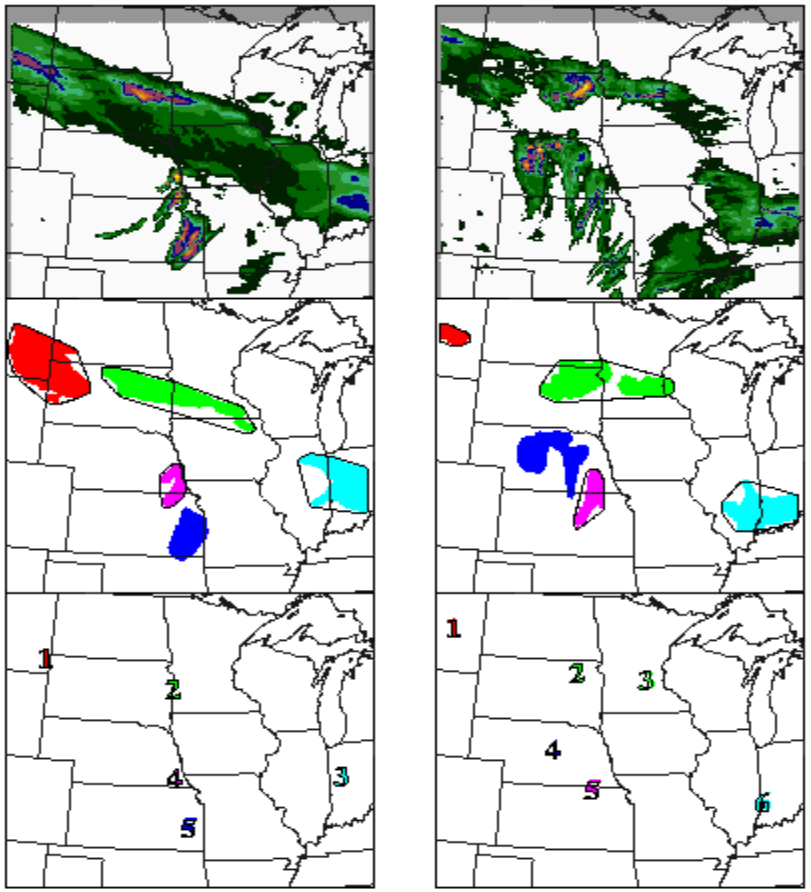
Displacement error	$0.142 in^2$	38.5%
Volume error	$0.000 in^2$	0.0%
Pattern error	$0.227 in^2$	61.5%
Conditional bias error	N.A.	N.A.


wrfnew 18-24h fcst from 20060406_00z run
Stage IV 6hr accum ending 20060406_00+24h
CRA valid 20060406_00+24h
Predicted rainfall (shifted)
Analyzed rainfall
wrfnew 18-24h fcst from 20060406_00z run
(42.82°,-107.04°) to (48.42°,-90.35°)
n=2353 CRA threshold=0.25 in/ 6 hr
Analysed Forecast
gridpoints ≥0.25 in/ 6 hr 748 1388
Average rainrate (in/ 6 hr) 0.62 0.61
Maximum rain (in/ 6 hr) 3.19 2.72
Rain volume (km³) 1.08 1.96
Displacement (E,N) = [0.80°,-0.86°]
Original Shifted
RMS error (in/ 6 hr) 0.57 0.55
Correlation coefficient -0.147 -0.102
Error Decomposition:
Displacement error 0.020in² 6.1%
Volume error 0.025in² 7.6%
Pattern error 0.282in² 86.2%
Conditional bias error N.A. N.A.


wrfnew 18-24h fcst from 20060406_00z run
Stage IV 6hr accum ending 20060406_00+24h
CRA valid 20060406_00+24h
Predicted rainfall (shifted)
Analyzed rainfall
wrfnew 18-24h fcst from 20060406_00z run
(43.29°,-107.04°) to (48.42°,-91.88°)
n=2084 CRA threshold=0.25 in/ 6 hr
Analysed Forecast
gridpoints ≥0.25 in/ 6 hr 748 1122
Average rainrate (in/ 6 hr) 0.62 0.94
Maximum rain (in/ 6 hr) 3.19 2.64
Rain volume (km³) 1.07 2.45
Displacement (E,N) = [-0.13°,-0.67°]
Original Shifted
RMS error (in/ 6 hr) 0.78 0.73
Correlation coefficient -0.075 0.068
Error Decomposition:
Displacement error 0.071in² 11.7%
Volume error 0.079in² 13.0%
Pattern error 0.457in² 75.3%
Conditional bias error N.A. N.A.

MODE: APCP_06 at SFC

Forecast | Observation

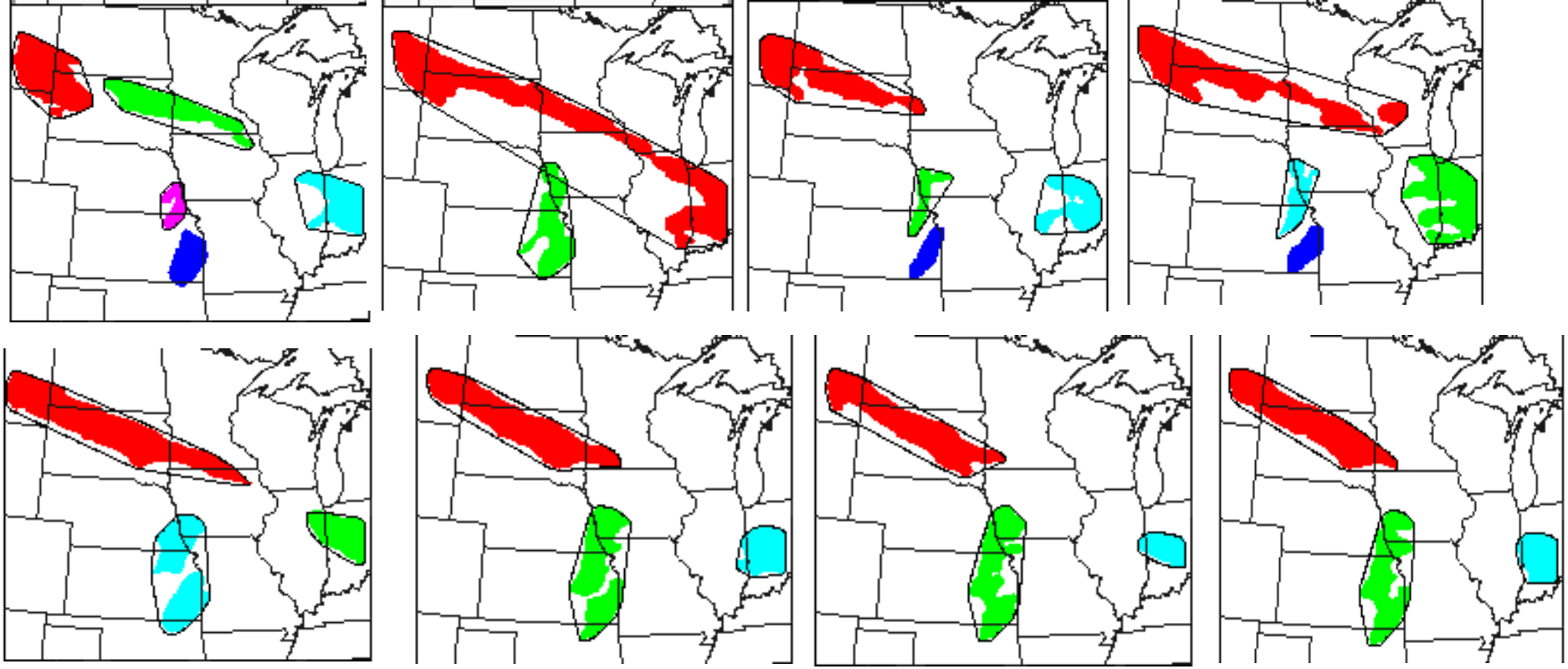


Fcst	Obs	Interest
4	5	0.9736
2	2	0.9254
3	6	0.9233
1	1	0.8798
2	3	0.7256
5	5	0.6837
4	4	0.6244
2	4	0.6110
5	4	0.5884
1	2	0.5652
1	3	N/A
4	1	N/A
3	3	N/A
4	3	N/A
5	3	N/A
1	4	N/A
5	1	N/A
3	4	N/A
2	1	N/A
3	1	N/A
1	5	N/A
2	5	N/A
3	5	N/A
3	2	N/A
4	2	N/A
1	6	N/A
2	6	N/A
5	2	N/A
4	6	N/A
5	6	N/A

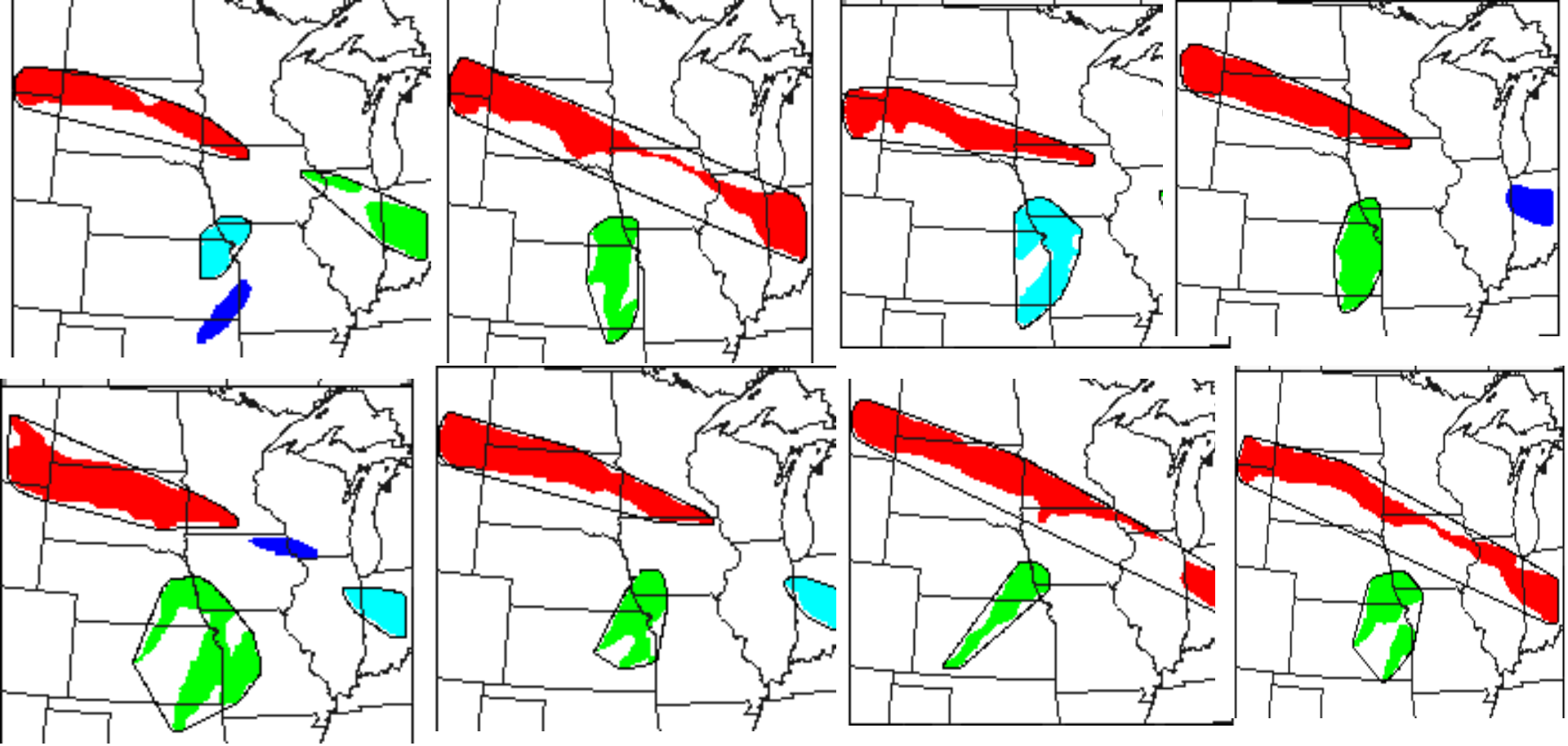
Init Time:	Apr 6, 2006 00:00:00
Valid Time:	Apr 7, 2006 00:00:00
Lead Time:	24 hours
Accum Time:	6 hours
Fuzzy Engine Weights	
Centroid Distance:	2.00
Boundary Distance:	4.00
Convex Hull Distance:	0.00
Angle Difference:	1.00
Area Ratio:	1.00
Intersection/Area:	2.00
Complexity Ratio:	0.00
Intensity Ratio:	0.00
Total Interest Thresh:	0.70

	Forecast	Observation
Mask Missing:	on	on
Mask Gr/Ply:	off/off	off/off
Raw Thresh:	<= 1000.00 mm	>= 0.00 mm
Conv Radius:	3 gs	3 gs
Conv Thresh:	>= 6.25 mm	>= 6.25 mm
Area Thresh:	>= 10 gs	>= 10 gs
Inten Perc:	102 th	102 th
Inten Thresh:	>= 1000.00 mm	>= 1000.00 mm
Merge Thresh:	>= 5.00 mm	>= 5.00 mm
Merging:	thresh	thresh
Matching:	match/fcst merge	
Simple(M/U):	5 (4/1)	6 (5/1)
Composites:	4	4

Example of MODE output for one ensemble member



April 6 18-24 hour forecast from Mixed Physics/Dynamics ensemble



Same forecast but from mixed initial/boundary condition ensemble

- **One question that can be explored is:**

  Does the behavior of spread/skill measures of object-based parameters match that of traditional measures applied to full model domains?

Spread measures in CRA and MODE

- For the parameters of:

  areal coverage,

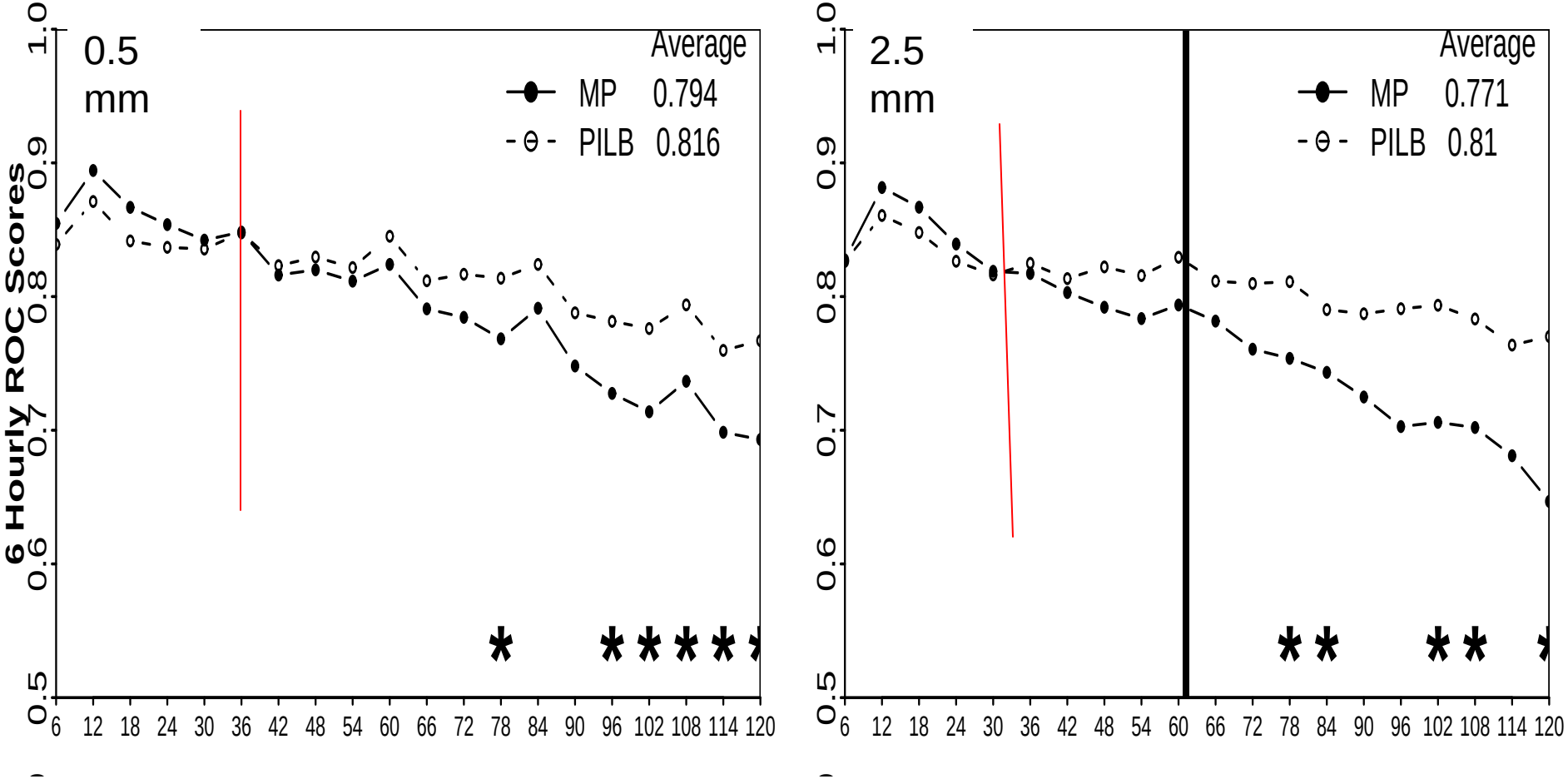
  mean rain rate,

  rain volume,

  number of systems,

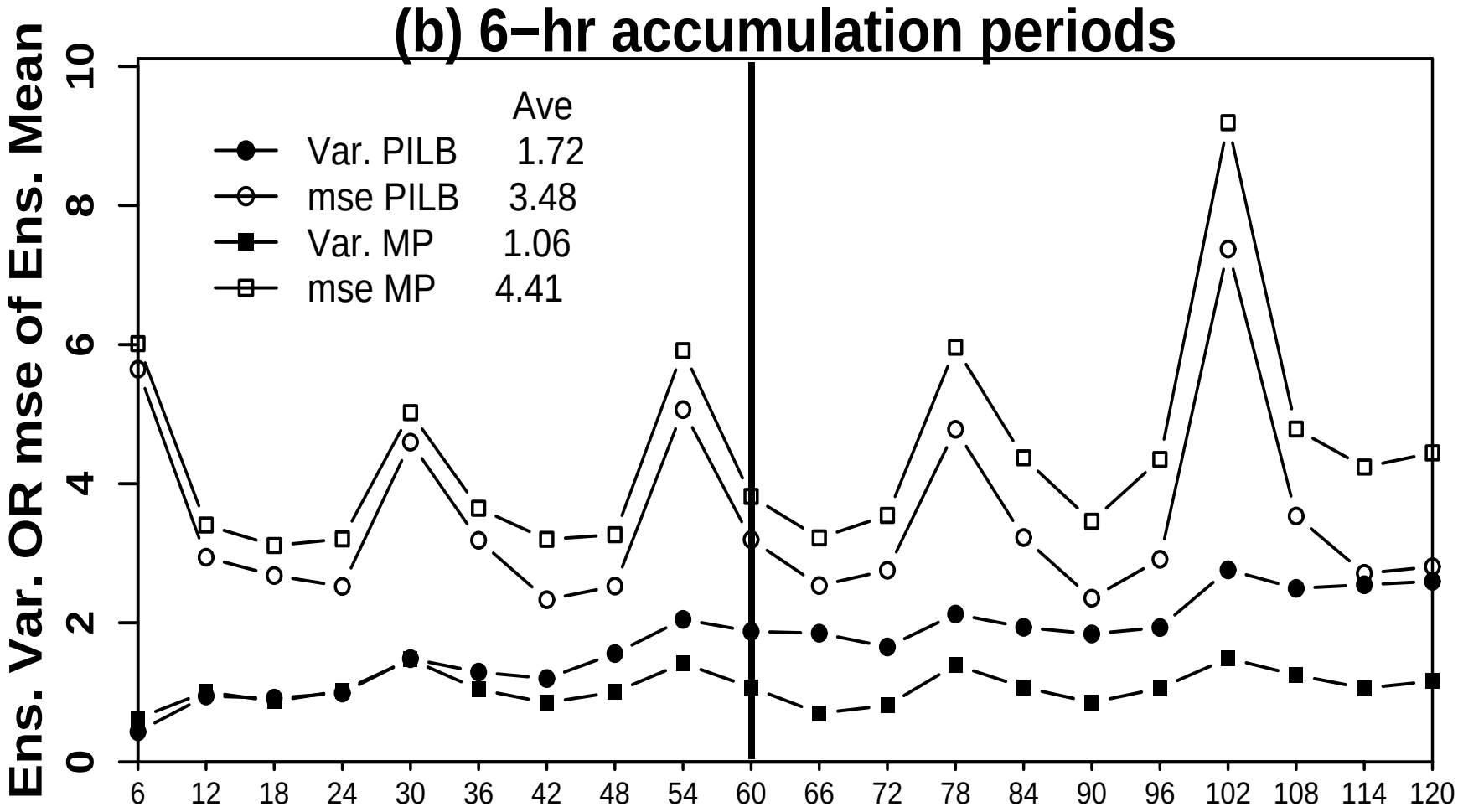
  and displacements

  standard deviations were computed as a function of time to compare trends to those found in Clark et al. (2008) for traditional ensemble spread measures.
- Clark et al. (2008) found spread & skill initially may have been best in mixed physics ensemble vs. IC/LBC one, but spread grew much faster in the IC/LBC one, and it performed noticeably better than mixed physics ensemble at later times (after 30-36 h) in these 120 h integrations.

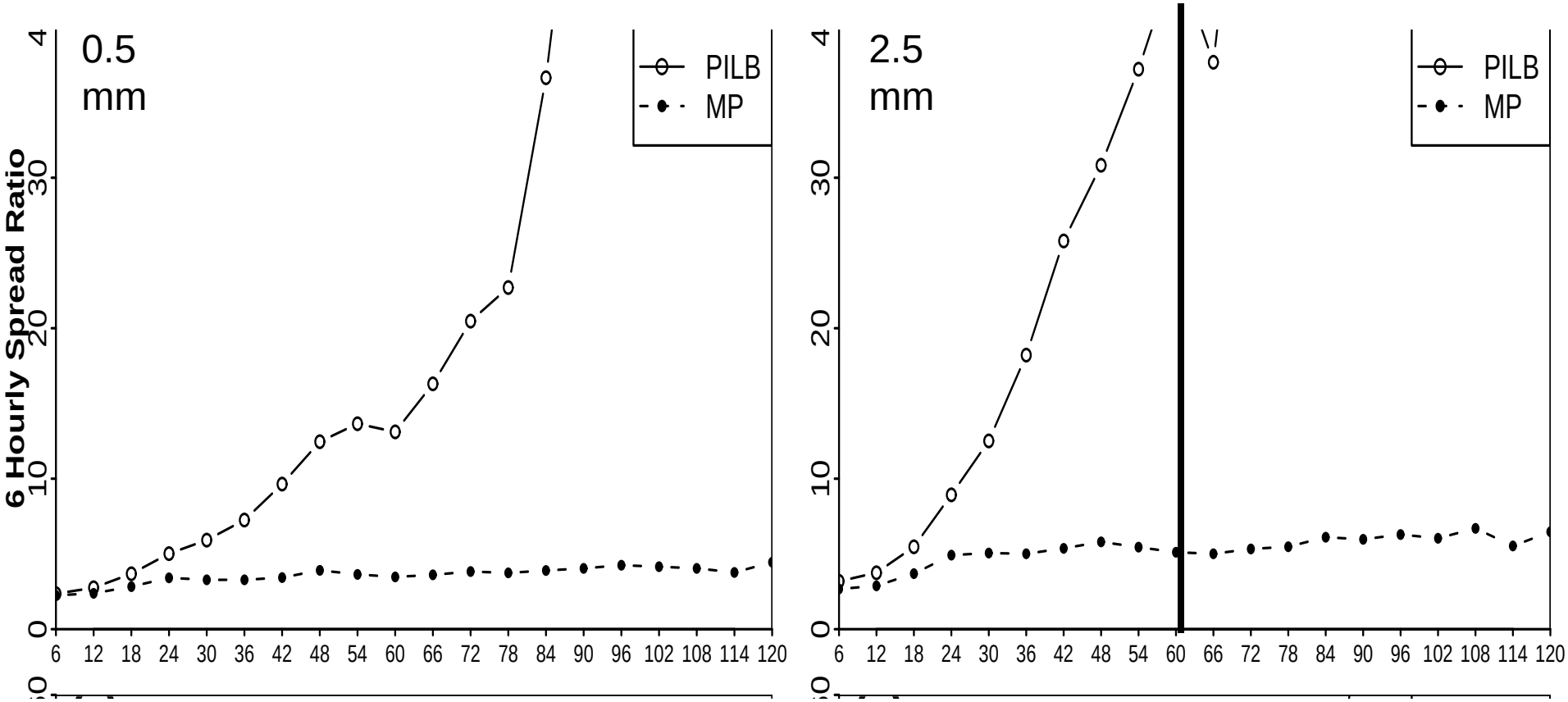


Areas under ROC curves for both ensembles (Clark et al. 2008)

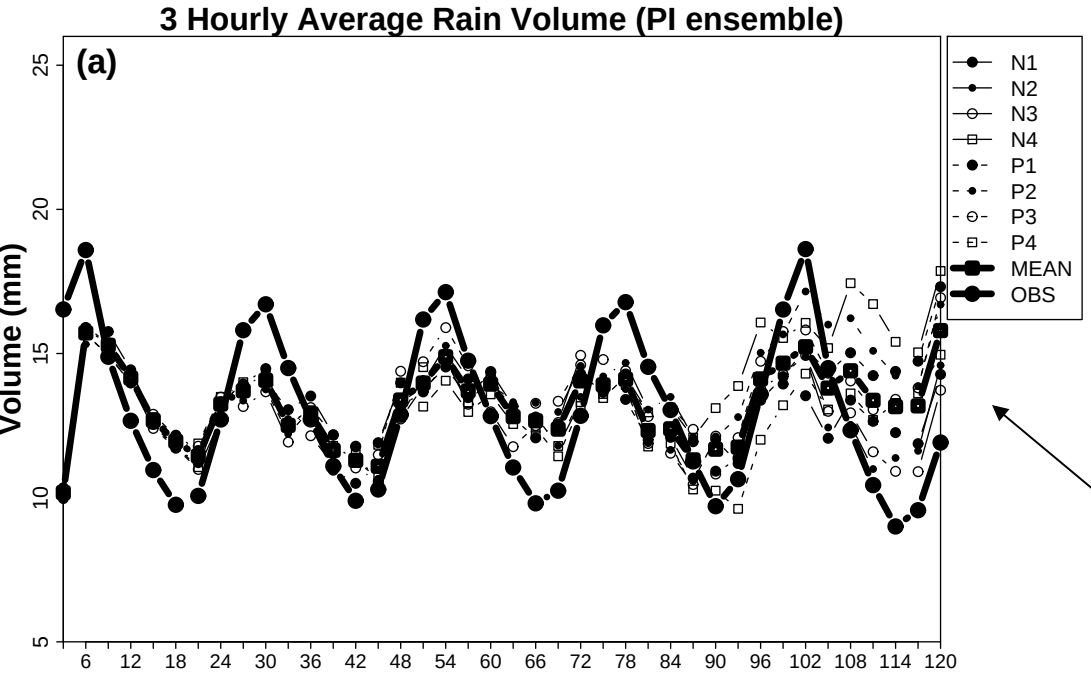
Skill initially better in mixed ensemble but IC/LBC becomes better after hour 30-36

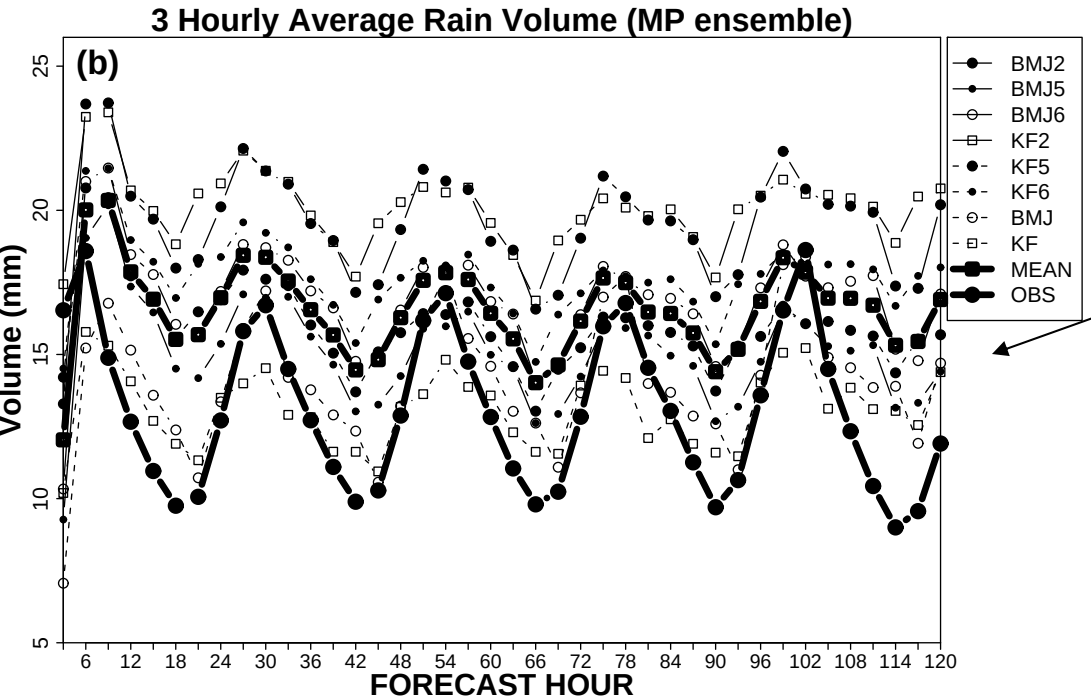


Variance continues to grow in IC/LBC ensemble but levels off after hour 30 in mixed ensemble. MSE always worse for mean of mixed ensemble – and performance worsens with time relative to IC/LBC ensemble.



Spread Ratio also shows dramatically different behavior with increasing spread in IC/LBC ensemble but little or no growth in mixed ensemble after first 24 hours





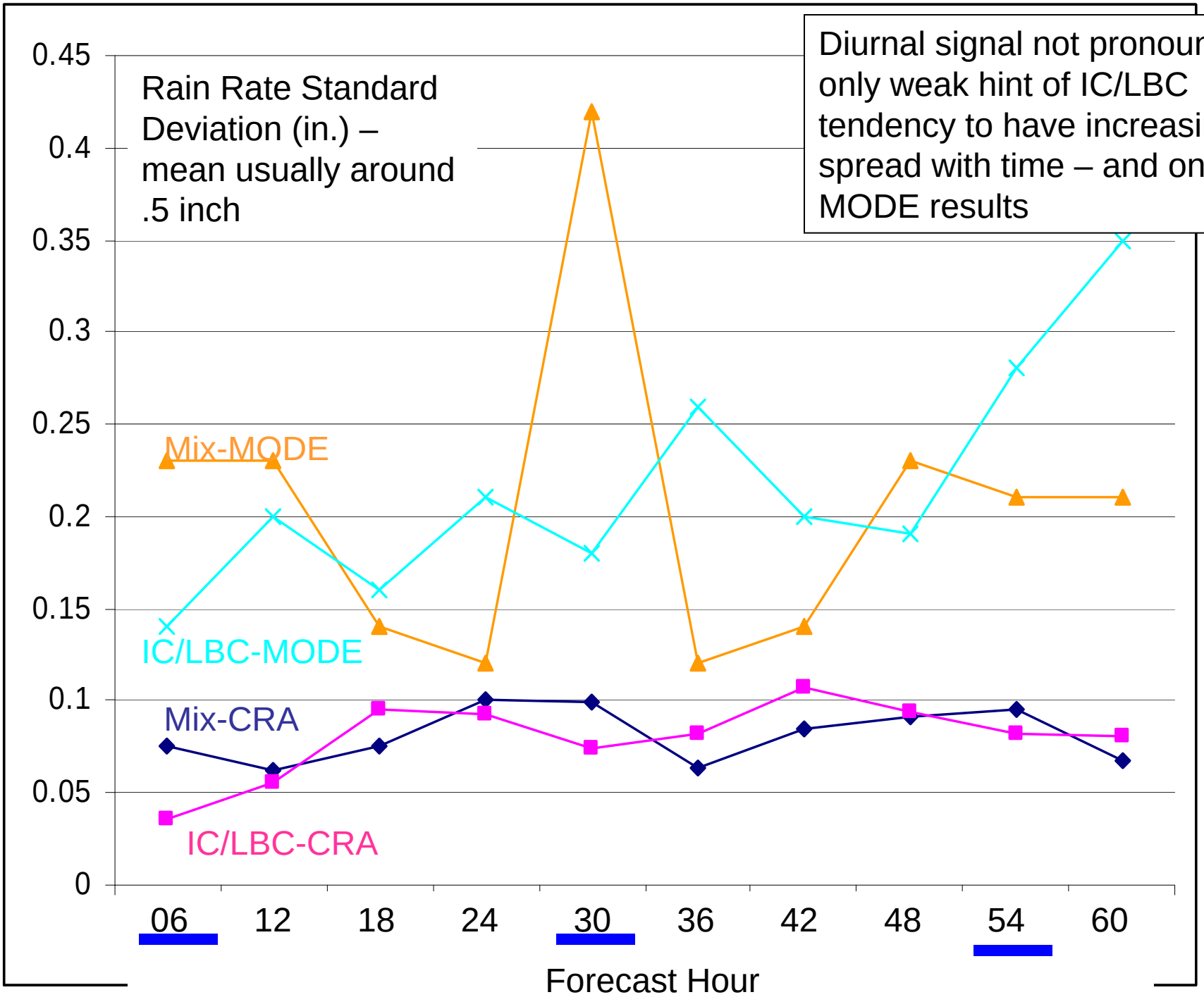
Observed Rain Volume (solid line) has peaks at hours 6, 30, 54, etc.

Amplitude of diurnal signal is reduced in IC/LBC ensemble (runs used NMM-KF-Ferrier)

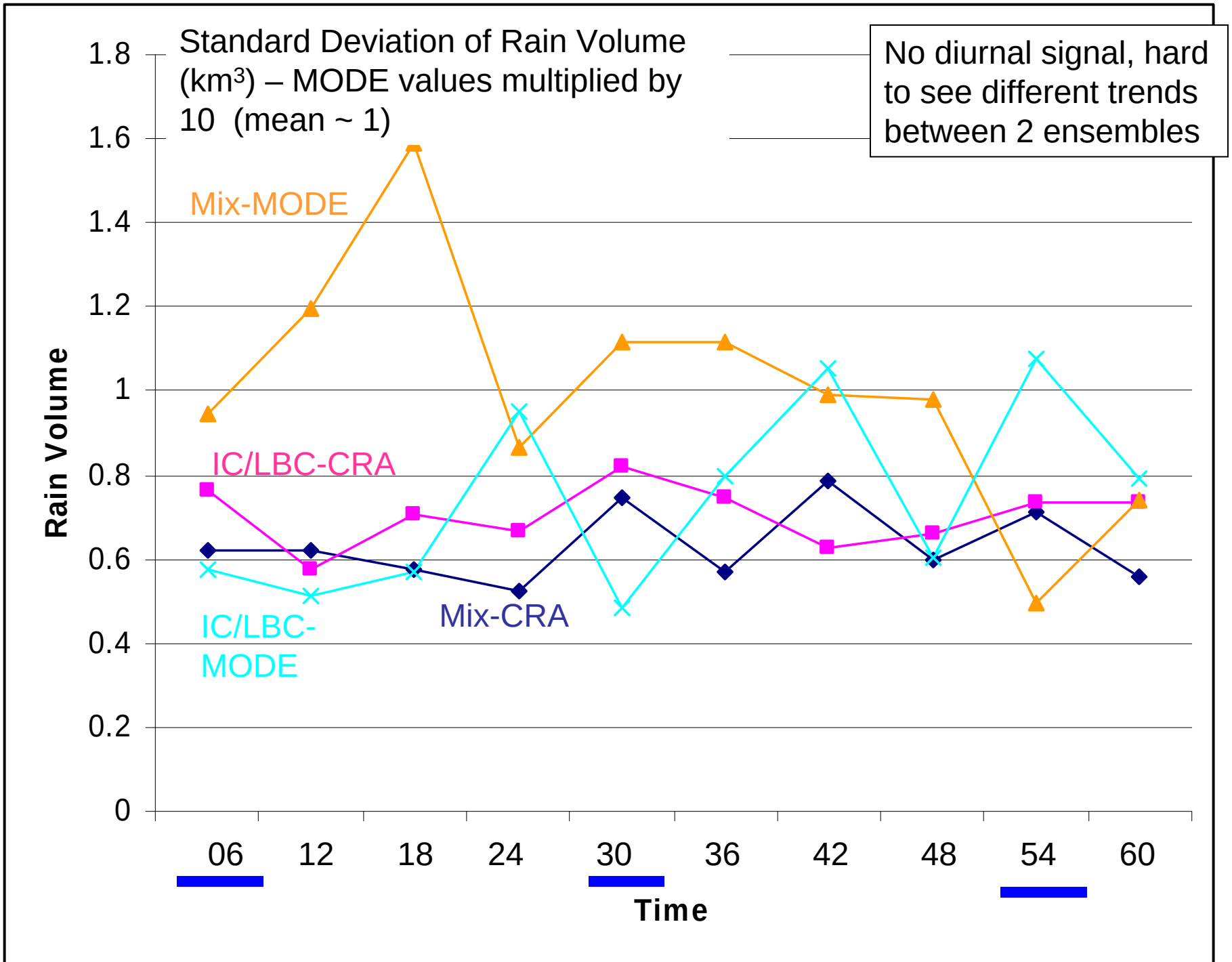
Volumes are overestimated by many of the mixed ensemble members

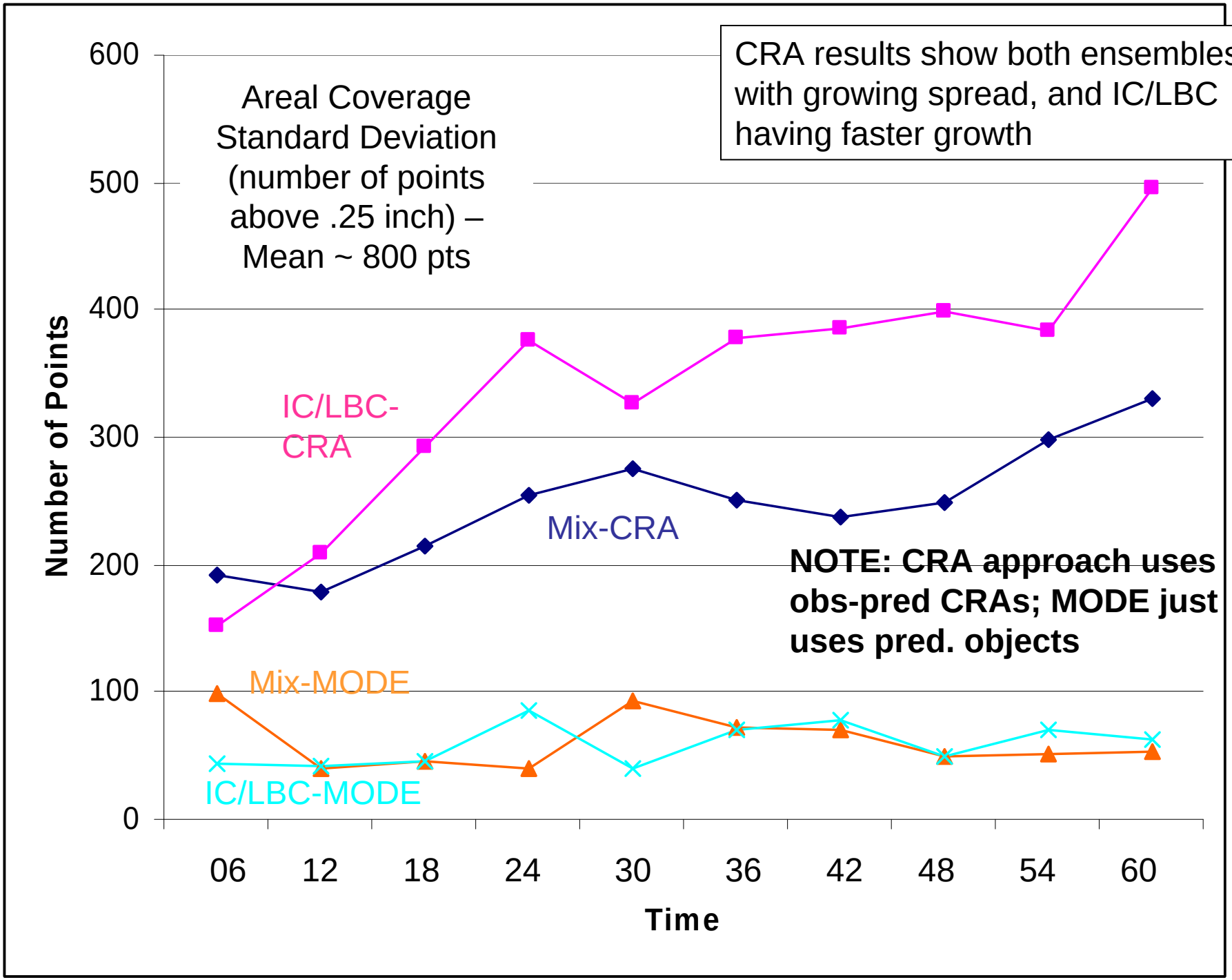
Questions:

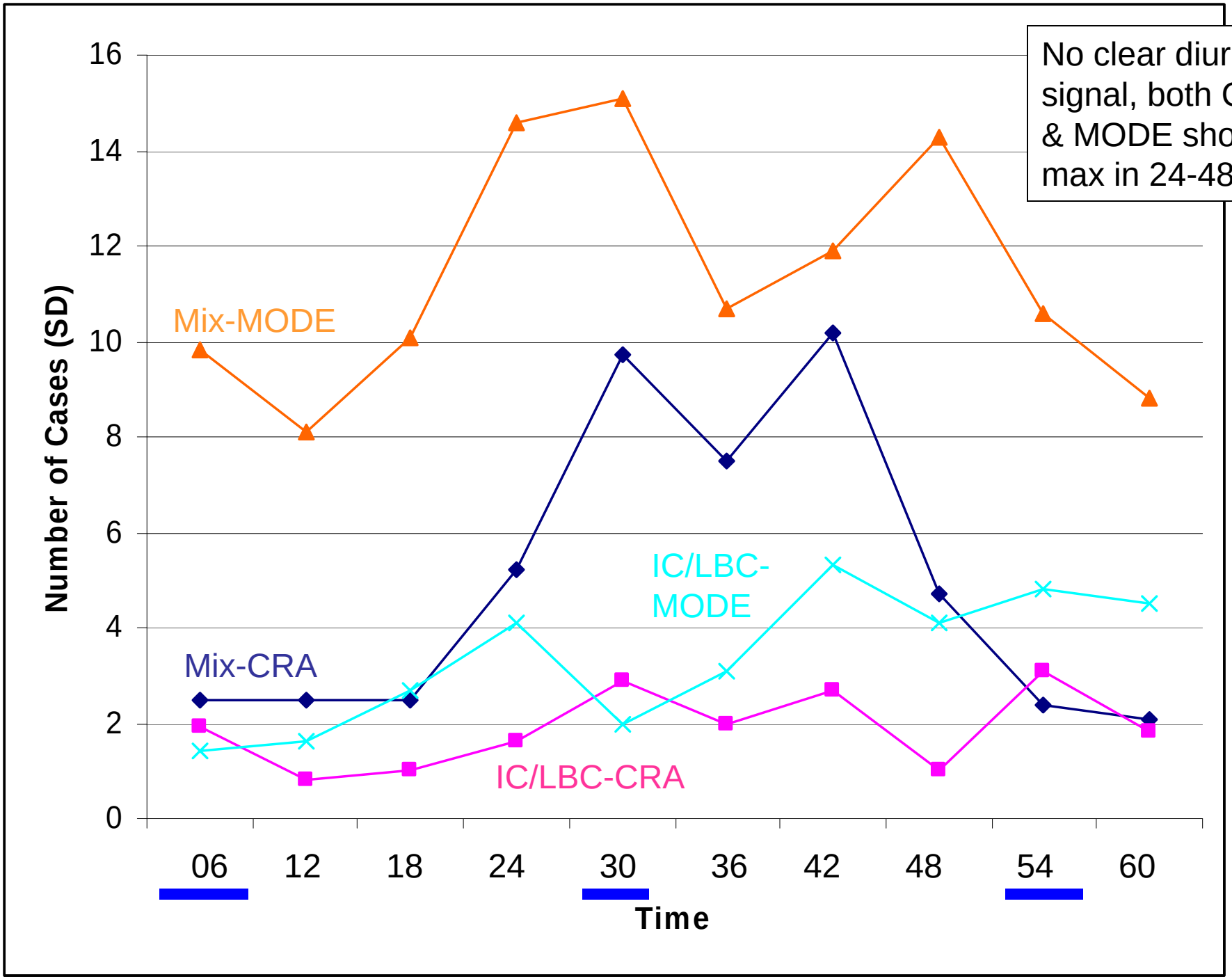
- Do the object parameters show the different behaviors between the Mix and IC/LBC ensembles?
- Do the object parameters show an influence from the diurnal trends in observed precipitation?

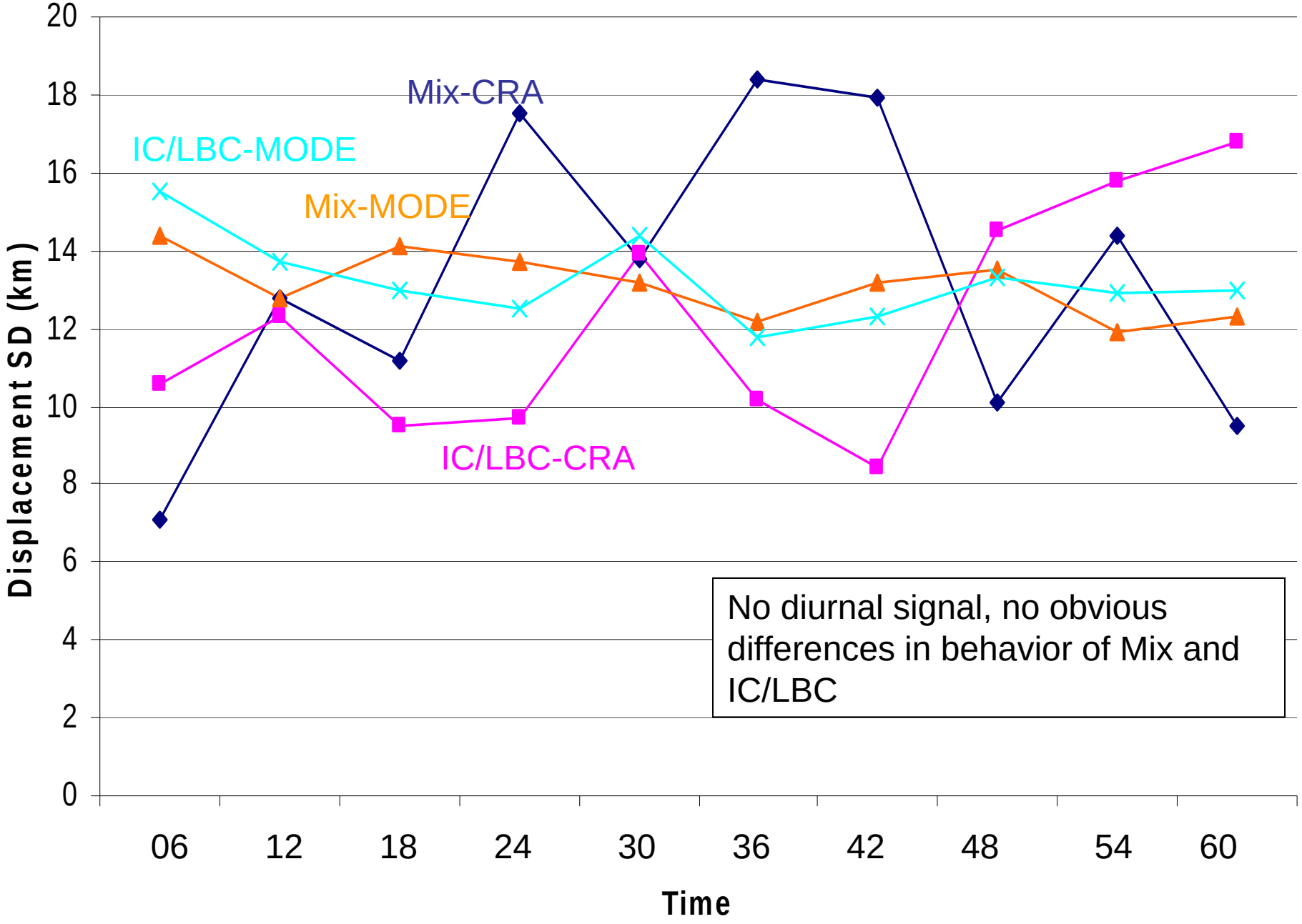


Wet times in blue


Standard Deviation of Rain Volume (km³) – MODE values multiplied by 10 (mean ~ 1)
No diurnal signal, hard to see different trends between 2 ensembles
Mix-MODE
IC/LBC-CRA
Mix-CRA
IC/LBC-MODE
Rain Volume
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
06
12
18
24
30
36
42
48
54
60
Time


CRA results show both ensembles with growing spread, and IC/LBC having faster growth
Areal Coverage Standard Deviation (number of points above .25 inch) – Mean ~ 800 pts
IC/LBC-CRA
Mix-CRA
Mix-MODE
IC/LBC-MODE
NOTE: CRA approach uses obs-pred CRAs; MODE just uses pred. objects
Number of Points
600
500
400
300
200
100
0
06
12
18
24
30
36
42
48
54
60
Time


No clear diurnal signal, both CRA & MODE show max in 24-48 h
Mix-MODE
Mix-CRA
IC/LBC-MODE
IC/LBC-CRA
Number of Cases (SD)
0
2
4
6
8
10
12
14
16
06
12
18
24
30
36
42
48
54
60
Time

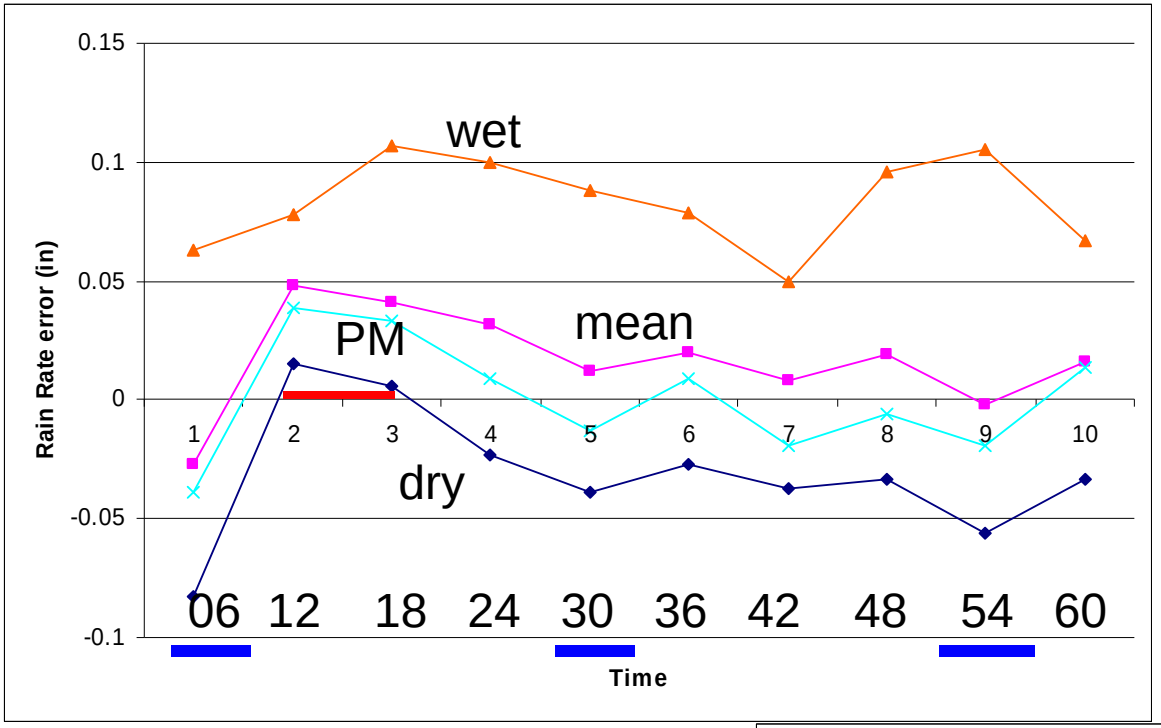


No diurnal signal, no obvious differences in behavior of Mix and IC/LBC

Other questions:

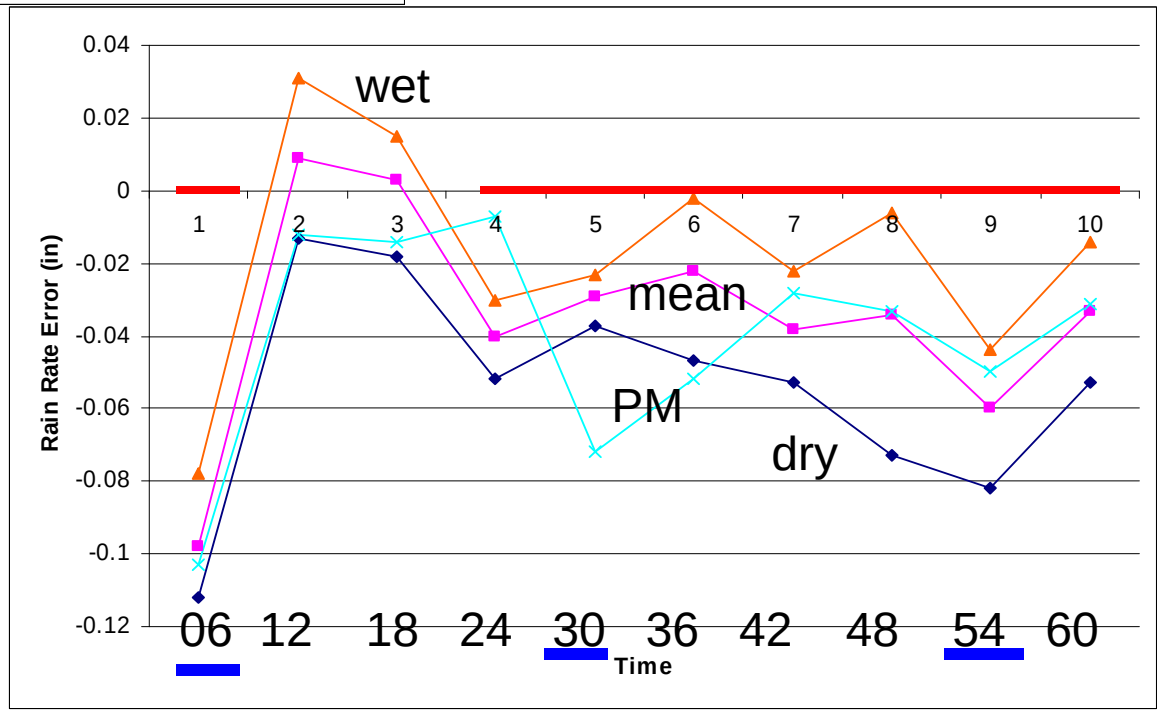
- Is the mean of the ensemble's distribution of object-based parameters a good forecast (better than those in an ensemble mean)?
- Does an increase in spread imply less predictability?
- How should a forecaster handle a case where only a subset of members show an object?

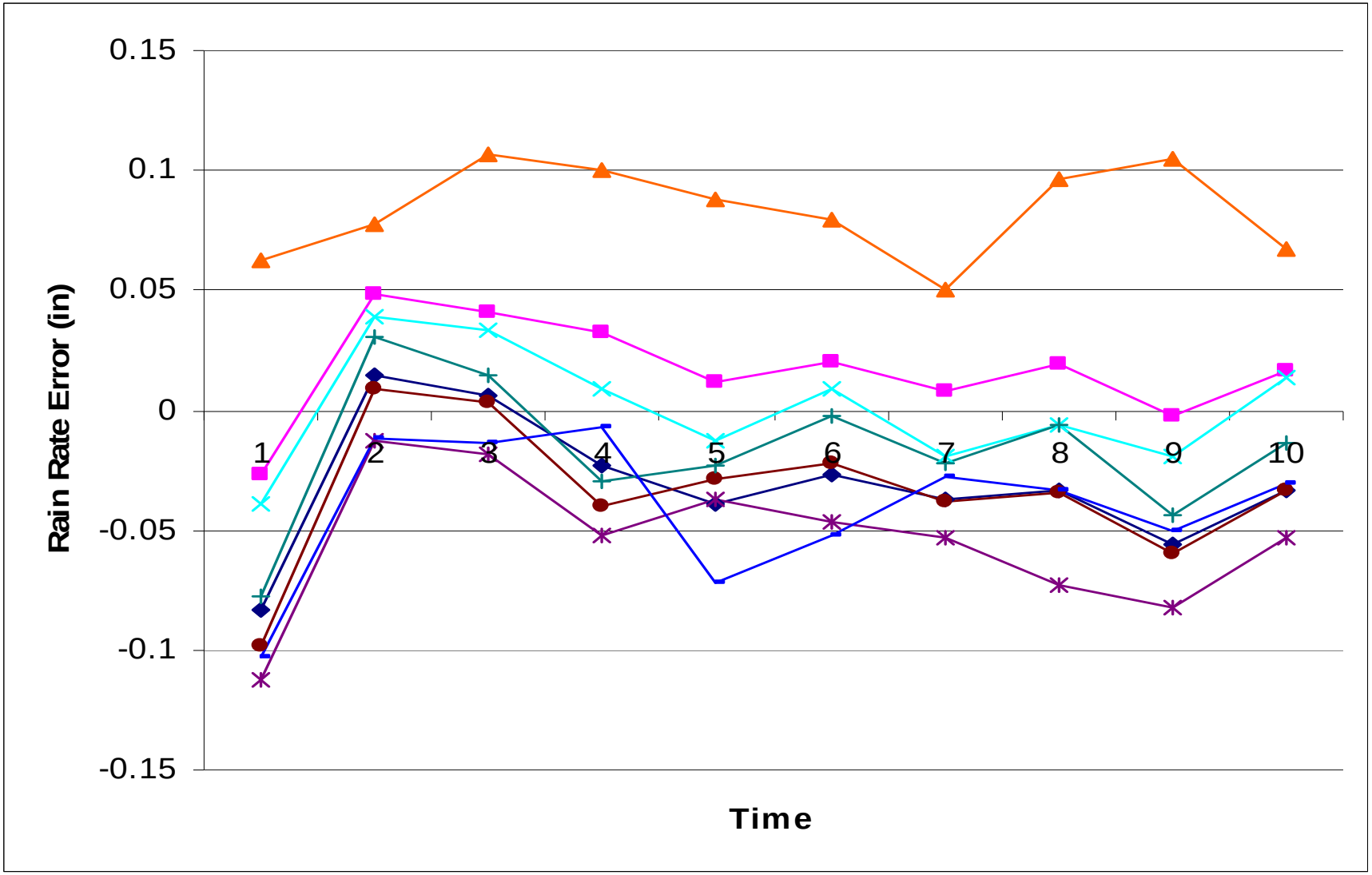
These questions have been examined using CRA



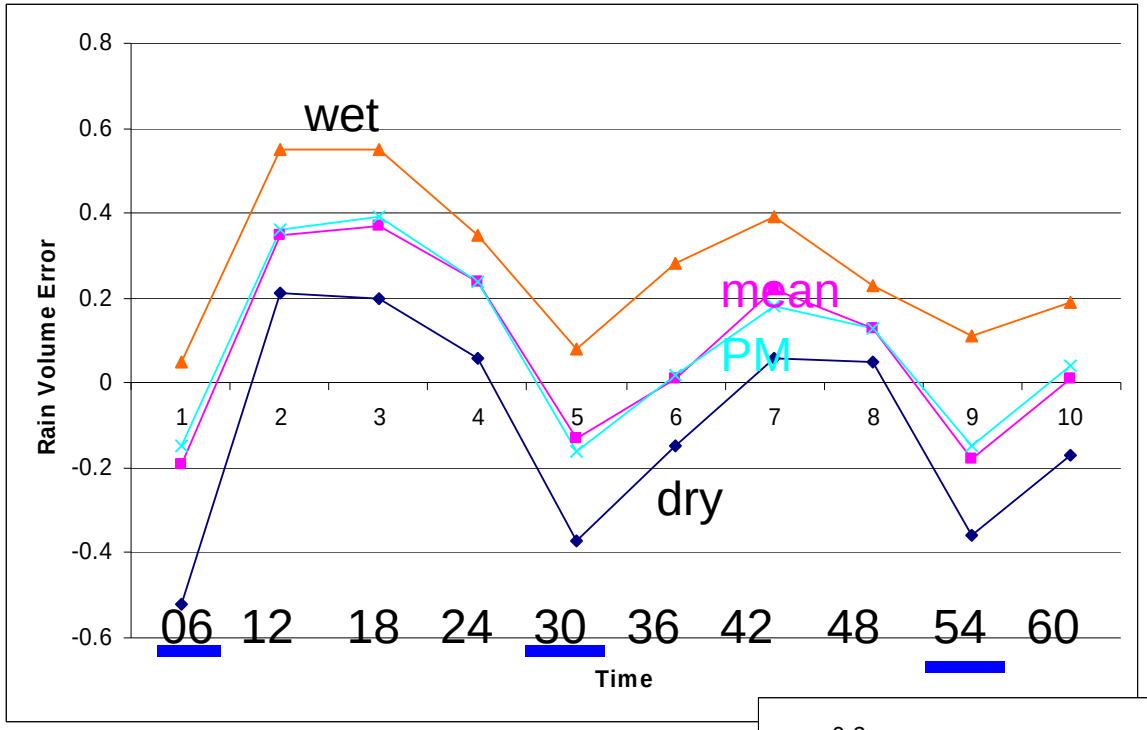
Mix Ensemble – in general, slight positive bias in rain rate, with Probability Matching forecast slightly less intense than mean of rates from members (PM usually better but not by much). Only during 06-18 period does observed rate not fall within forecasted range.

IC/LBC – usually too dry with rain rate (at all hours except 06-18), Probability Matching forecast exhibits much more variable behavior, again its performance is comparable to mean of rates of members

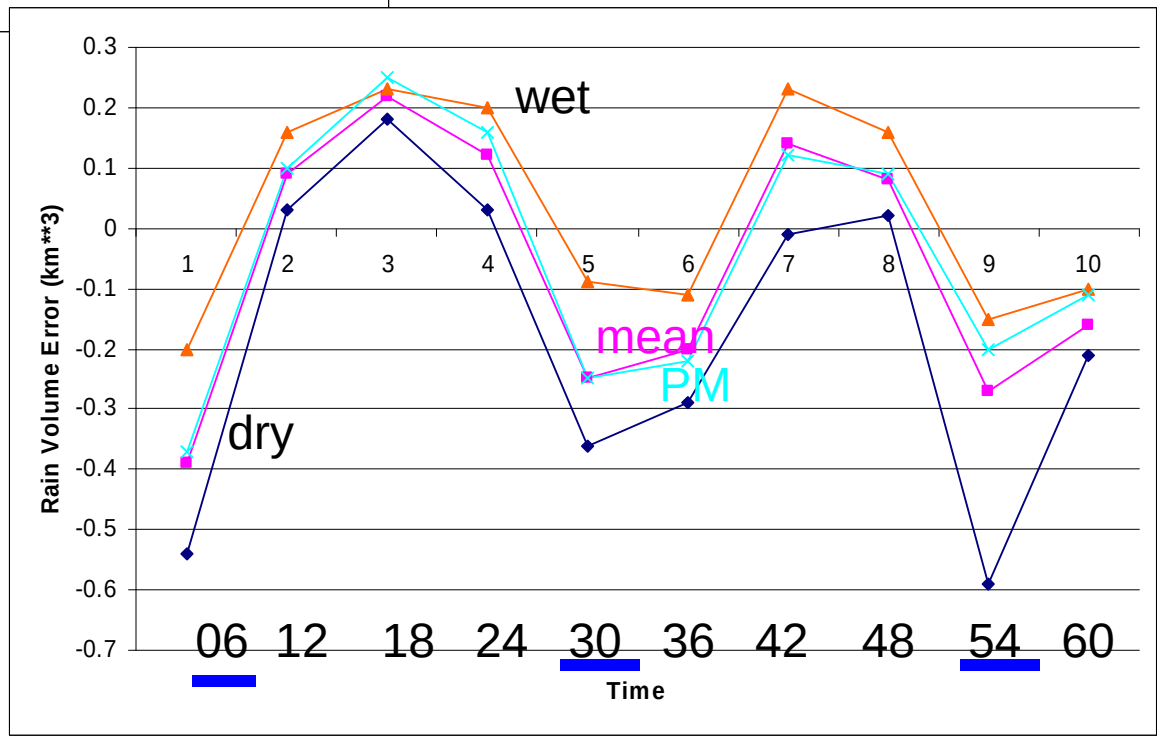




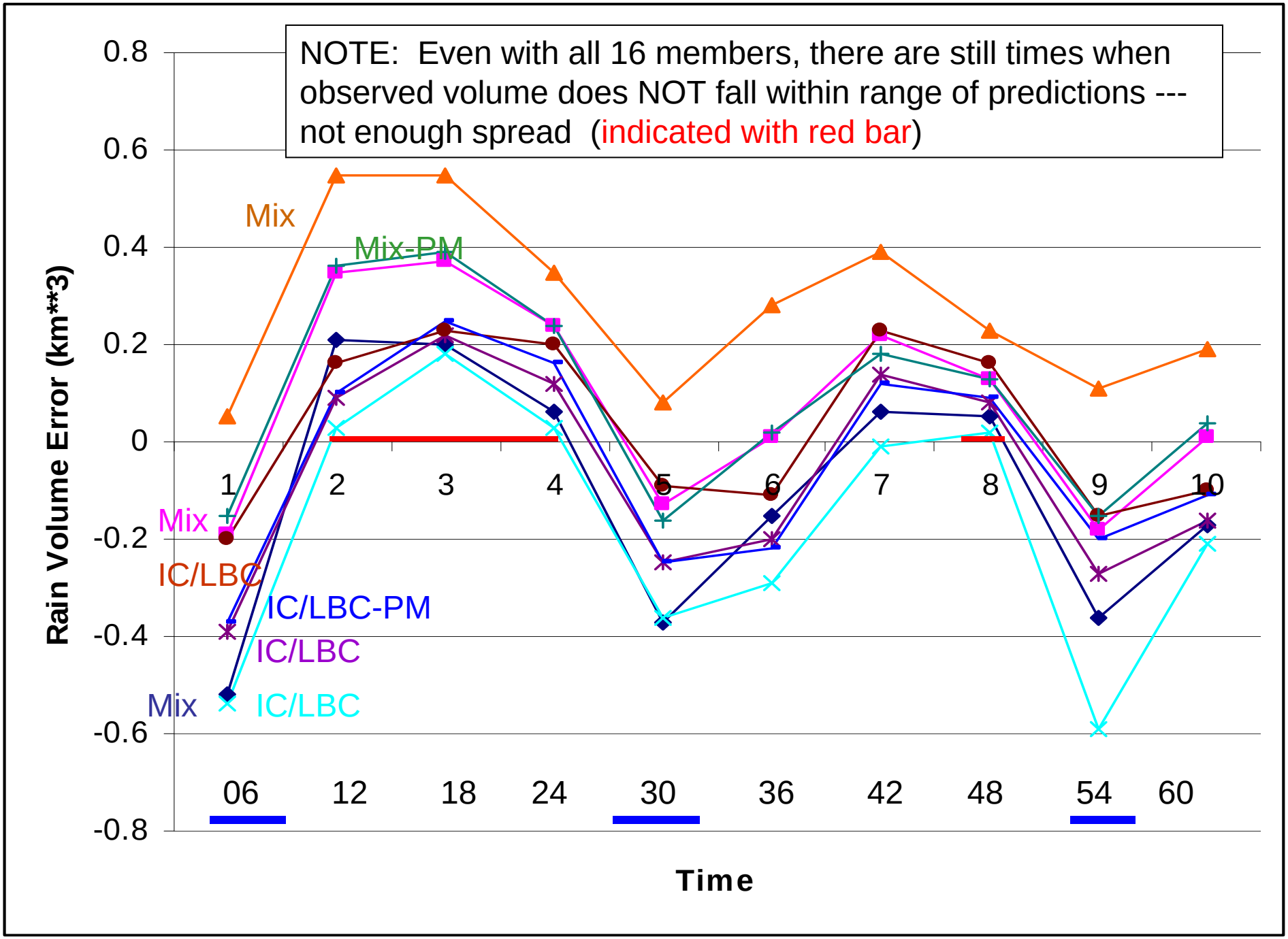
Notice that at all times, the observed rain rate falls within the range of values from the full 16 member ensemble – indicating potential value for forecasting



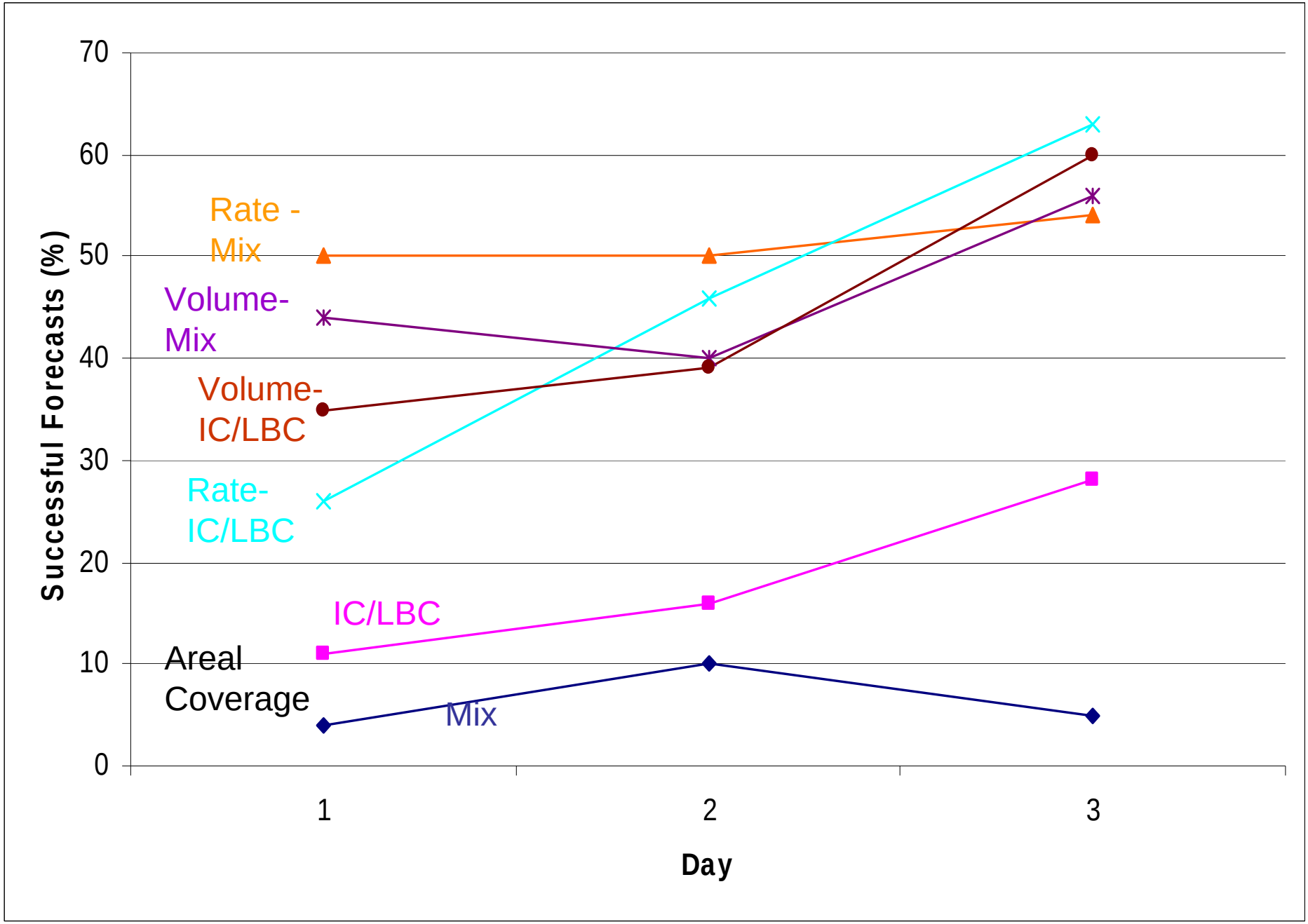
Mix Ensemble – clear diurnal signal, usually too much rain volume except at times of observed peak, when it is too small. Probability Matching equal in skill to mean of member volumes

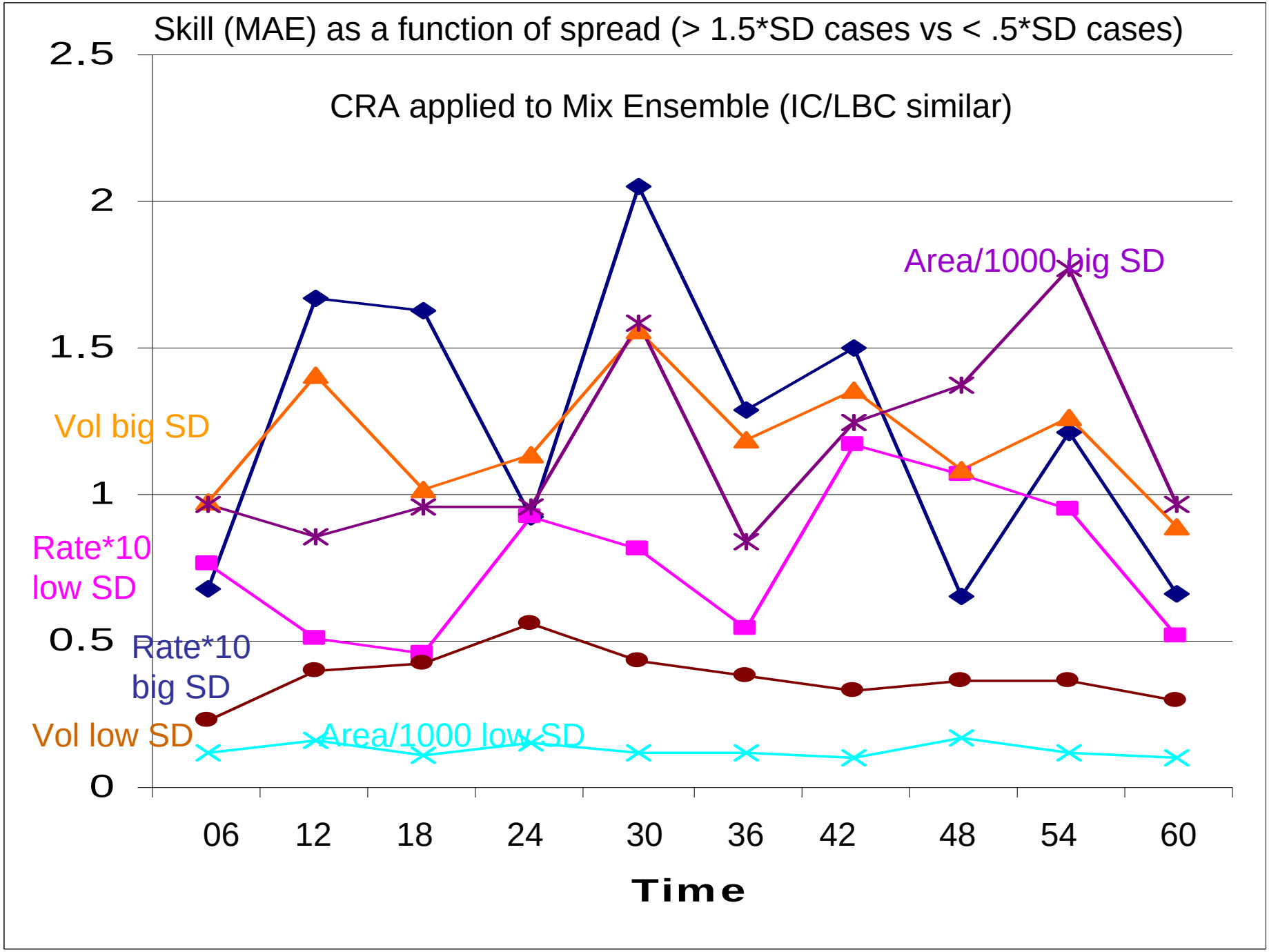


IC/LBC Ensemble – also clear diurnal signal, less volume than Mix ensemble, Probability Matching usually a little wetter but generally comparable to mean of members


NOTE: Even with all 16 members, there are still times when observed volume does NOT fall within range of predictions --- not enough spread (indicated with red bar)
Rain Volume Error (km**3)
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
Mix
Mix-PM
Mix
IC/LBC
IC/LBC-PM
IC/LBC
Mix
IC/LBC
1
2
3
4
5
6
7
8
9
10
06
12
18
24
30
36
42
48
54
60
Time

Percentage of times the observed value fell within the min/max of the ensemble




Skill (MAE) as a function of spread (> 1.5*SD cases vs < .5*SD cases)
CRA applied to Mix Ensemble (IC/LBC similar)
2.5
2
1.5
1
0.5
0
Area/1000 big SD
Vol big SD
Rate*10
low SD
Rate*10
big SD
Vol low SD
Area/1000 low SD
06
12
18
24
30
36
42
48
54
60
Time

- **It thus appears that total system rain volume and total system areal coverage of rainfall show a clear signal for better skill when spread is smaller**
- **Rain rate does not show such a clear signal – (especially when 4 bins of SDs are examined).**
- ***Perhaps average rain rate for systems is not as big a problem in the forecasts as areal coverage (and thus volume)? Seems to be ~5-10% error for rate, 10-20% for volume, 10-20% for area***

Summary

- Ensemble spread behavior for object-oriented parameters may not behave like traditional ensemble measures
- Some similarities but some differences also in output from CRA vs MODE
- Some suggestion that ensembles may give useful information on probability of systems having a particular size, intensity, volume

Future Work

- How can location information best be determined from ensembles?
- Can lack of spread in the object-oriented parameters be improved entirely with addition of reasonable number of members?
- How well does a forecast of object parameters based on probability information perform?
- How will results differ based on grid spacing -- 4 km vs 20 km ensemble output will be explored

Acknowledgments

- Funding provided by WRF-DTC and through NSF grant ATM-0537043
- Thanks to Adam Clark for providing the precipitation output
- Thanks to Randy Bullock and John Halley-Gotway for MODE and R help at NCAR
- Thanks to Eric Aligo and Daryl Herzmann for assistance with Excel and computer issues