A few slides on forecast bias correction

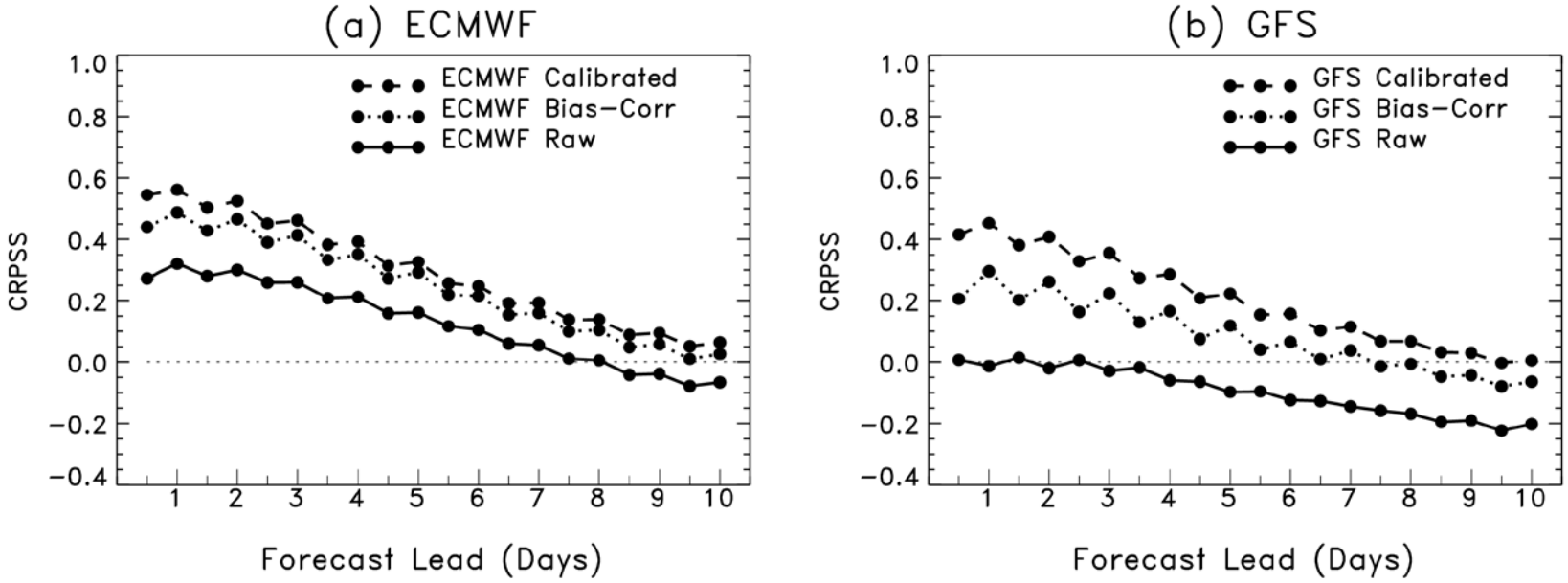
Matt Newman, University of Colorado/CIRES and NOAA/ESRL/PSD

Bias corrections

- Mean error
  - forecast error based on forecast lead time alone
- State-dependent error
  - systematic mean errors for different forecast states
- Correction of forecast ensemble spread
  - Models typically are too “certain” (ensemble has too-small spread)

Gross bias correction

- Given sample of past forecasts $x_1, \ldots, x_n$ and observations $y_1, \ldots, y_n$, gross bias correction is simply $\overline{y} - \overline{x}$

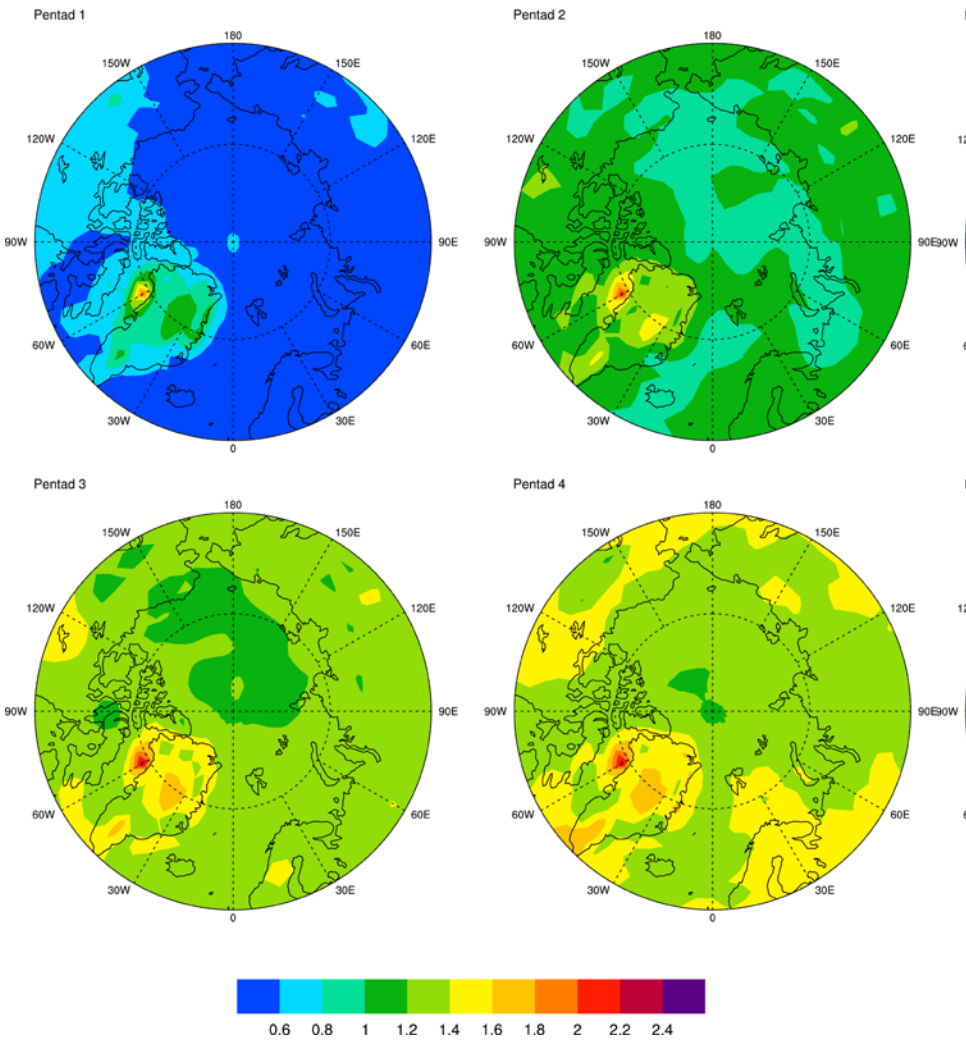


In surface-temperature calibration experiments with NCEP's GFS and ECMWF, simple gross bias correction achieved a large percentage of the improvement that was achieved through more sophisticated, bias+spread correction.

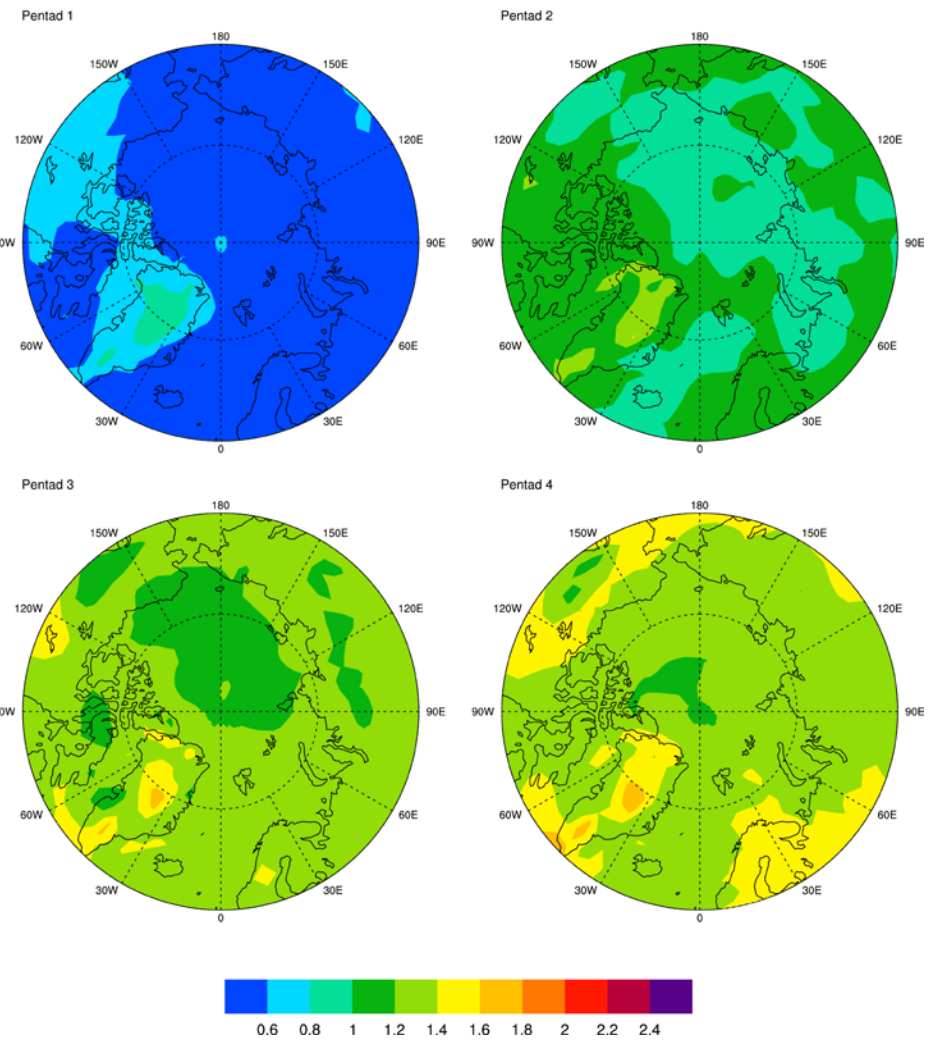
Ref: Hagedorn et al., MWR, 2008.

HiCOM-FIM 850mb zonal wind hindcast skill (rms forecast error)

Error without bias correction



Error with bias correction



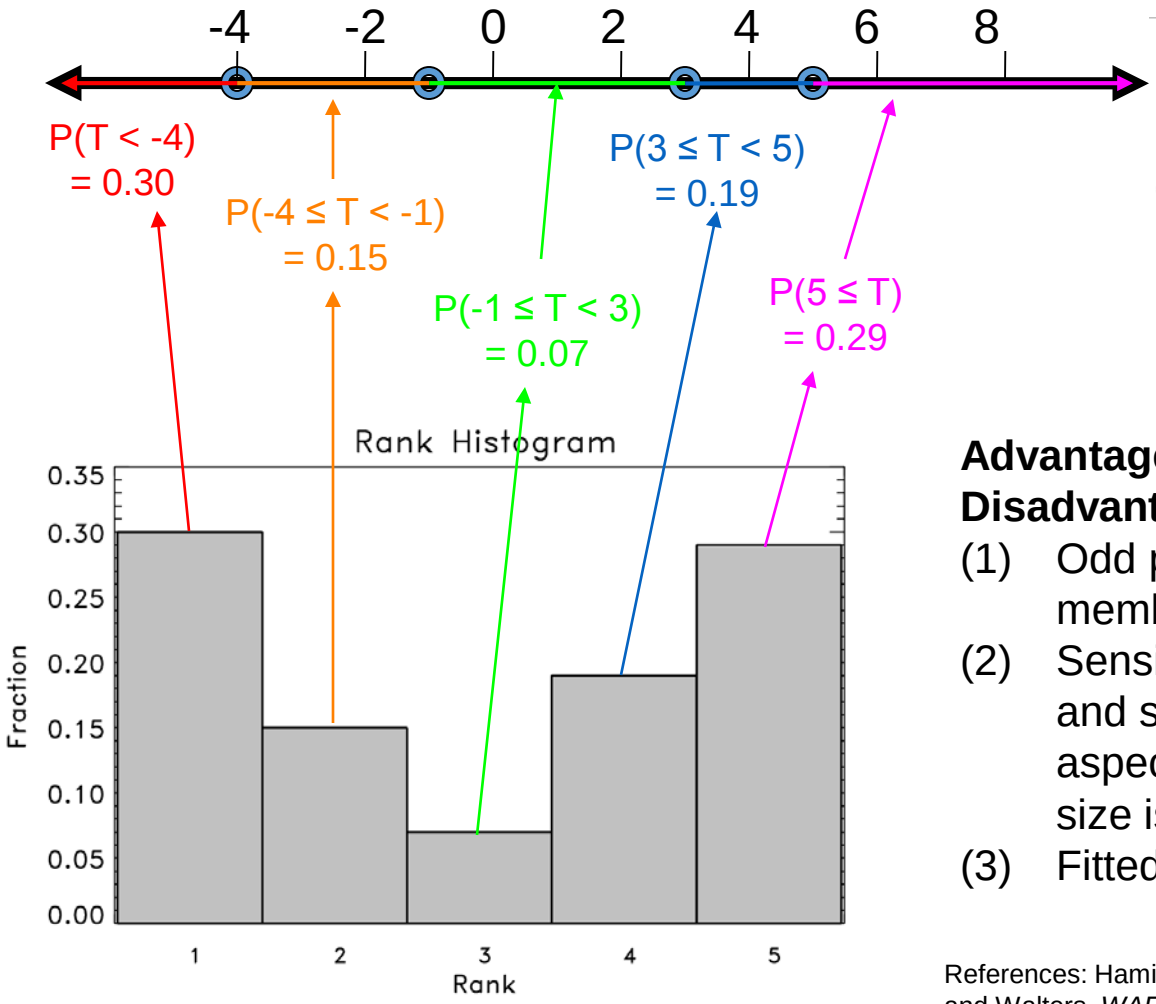
Model Output Statistics ("MOS")

many elements based on multiple linear regression

```
KBID   GFS MOS GUIDANCE    2/16/2005  1800 UTC
 DT /FEB  17                   /FEB  18                /FEB  19
 HR   00 03 06 09 12 15 18 21 00 03 06 09 12 15 18 21 00 03 06 12 18
 N/X              32             40             25             35       19
 TMP  42 39 36 33 32 36 38 37 35 33 30 28 27 30 32 31 28 25 23 19 27
 DPT  34 29 26 22 19 18 17 17 17 17 17 15 14 13 11  8  7  6  5  2  4
 CLD  OV FW CL CL SC BK BK BK BK BK BK BK SC BK BK BK BK FW CL CL CL
 WDR  26 30 32 32 32 31 29 28 30 32 31 31 31 31 30 29 31 32 33 33 27
 WSP  12 12 12 11 08 08 09 08 09 09 10 10 10 12 13 13 15 16 15 09 08
 P06        17      0      0      0      4      0     10      6      8  0  0
 P12               17             0             10             17       8
 Q06          0      0      0      0      0      0      0      0      0  0  0
 Q12               0             0             0              0        0
 T06       0/ 2   0/ 0   1/ 0   1/ 2   0/ 1   0/ 1   1/ 0   0/ 1   0/ 0   0/ 0
 T12                      1/ 0          1/ 2          1/ 1          0/ 1   0/ 0
 POZ   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
 POS  13  47  70  84 91100 96100100100100 92100 98100100100 94 92100100
 TYP   R   S   S   S   S   S   S   S   S   S   S   S   S   S   S   S   S   S   S   S
 SNW                                                  0                       0
 CIG   7   8   8   8   8   8   8   8   8   7   7   7   8   7   7   7   8   8   8   8   8
 VIS   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7   7
 OBV   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N   N
```

US: Statistical corrections to operational US NWS models, some fixed (NGM), some not (Eta, GFS). Refs: http://www.nws.noaa.gov/mdl/synop/index.htm, Carter et al., *WAF*, **4**, p 401, Glahn and Lowry, *JAM*, **11**, p 1580. **Canadian** models discussed in Wilson and Vallee, *WAF*, **17**, p. 206, and *WAF*, **18**, p 288. **Britain**: Met Office uses "updateable MOS" much like perfect prog.

Ensemble calibration: rank histogram techniques



NCEP MRF precipitation forecasts, from Eckel and Walters, 1998

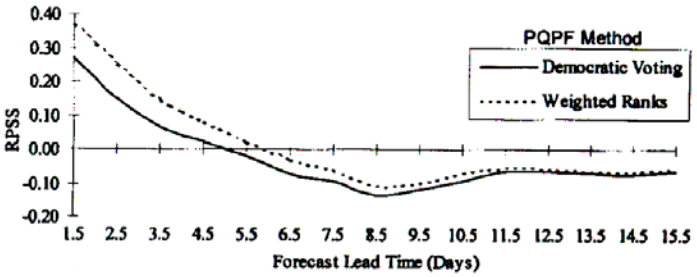


FIG. 10. Ranked probability skill score (RPSS) results for all forecast lead times.

Advantages: Demonstrated skill gain

Disadvantages:

(1) Odd pdfs, especially when two ensemble members close in value.

(2) Sensitive to shape of rank histogram, and shape of histogram may vary with aspects like precip amount --> sample size issues.

(3) Fitted parametric distributions as skillful

References: Hamill and Colucci (*MWR*, 1997, 1998; Eckel and Walters, *WAF*, 1998; used at Met Office)

Questions to consider

- How many samples do we need to determine bias?
- How much data/hindcasts do we actually have?
- What kind of bias do we need to remove?