Deconstruction of a science paper’s data-evidence basis

Xingchen Yang

MPO 624

Jan 27, 2018

My Choice

- Title, citation
  - ➢ Probabilistic Quantitative Precipitation Forecasting Using Bayesian Model Averaging
  - ❖ J. Mclean Sloughter, Adrian E. Raftery, Tilmann Gneiting, and Chris Fraley, Department of Statistics, University of Washington, Seattle, Washington

- Size of evidence sets
  - o 8 figures, 3 tables
  - o 1 magic number (perhaps)

Feature is claimed to exist in Fig.1

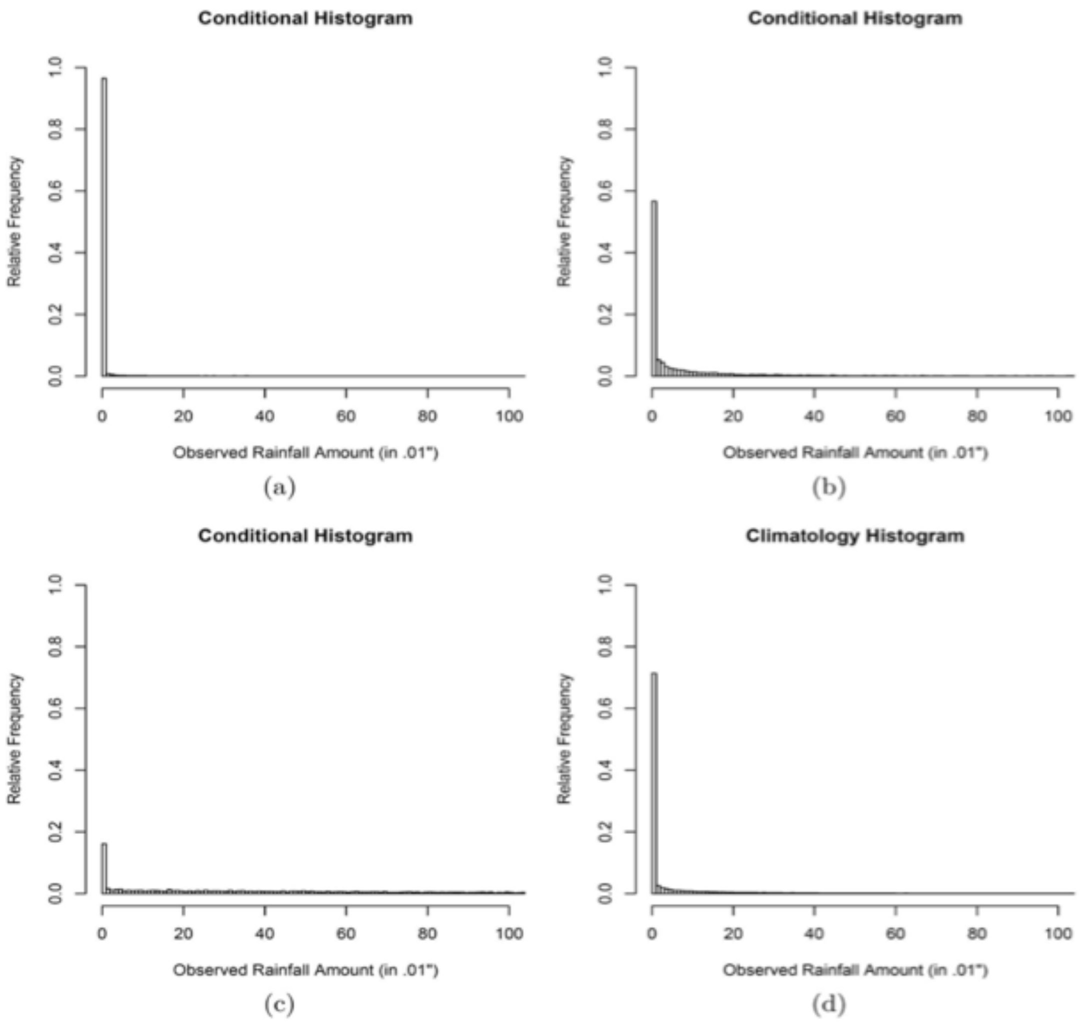
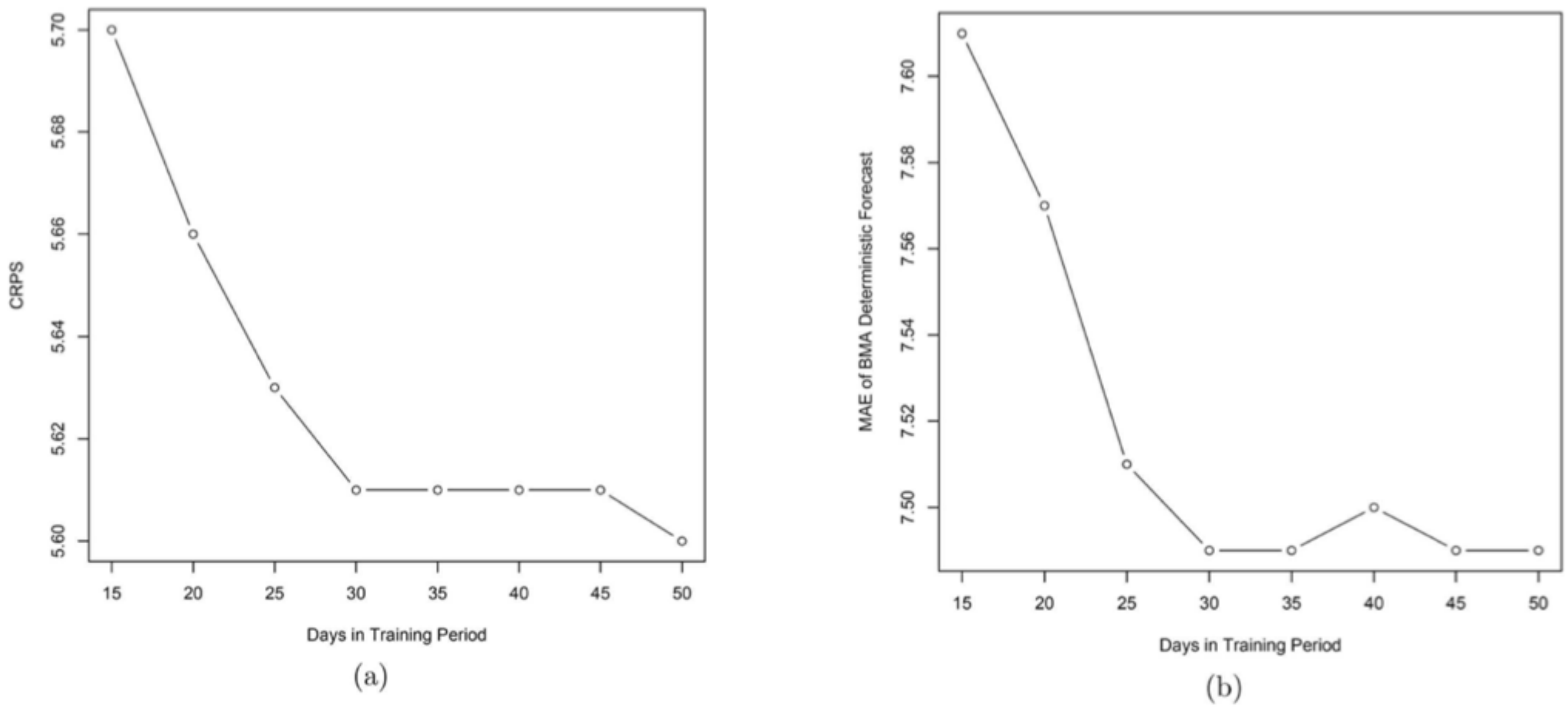


Fig.1, Histograms of observed precipitation accumulation for cases in which the centroid member forecast of precipitation was (a) zero, (b) between 6.4 and 9.6 hundredths of an inch, and (c) greater than 0.0594 in., and (d) all cases.

Gamma distribution was fit to the precipitation amount. (this paper uses cube root of the precipitation)

Optimal model parameters found in Fig. 2



Optimal training days N=30 is determined in Fig. 2

FIG. 2. Comparison of training period lengths: (a) CRPS of BMA forecasts and (b) MAE of BMA deterministic forecasts.

Illustration of method shown in Fig. 3

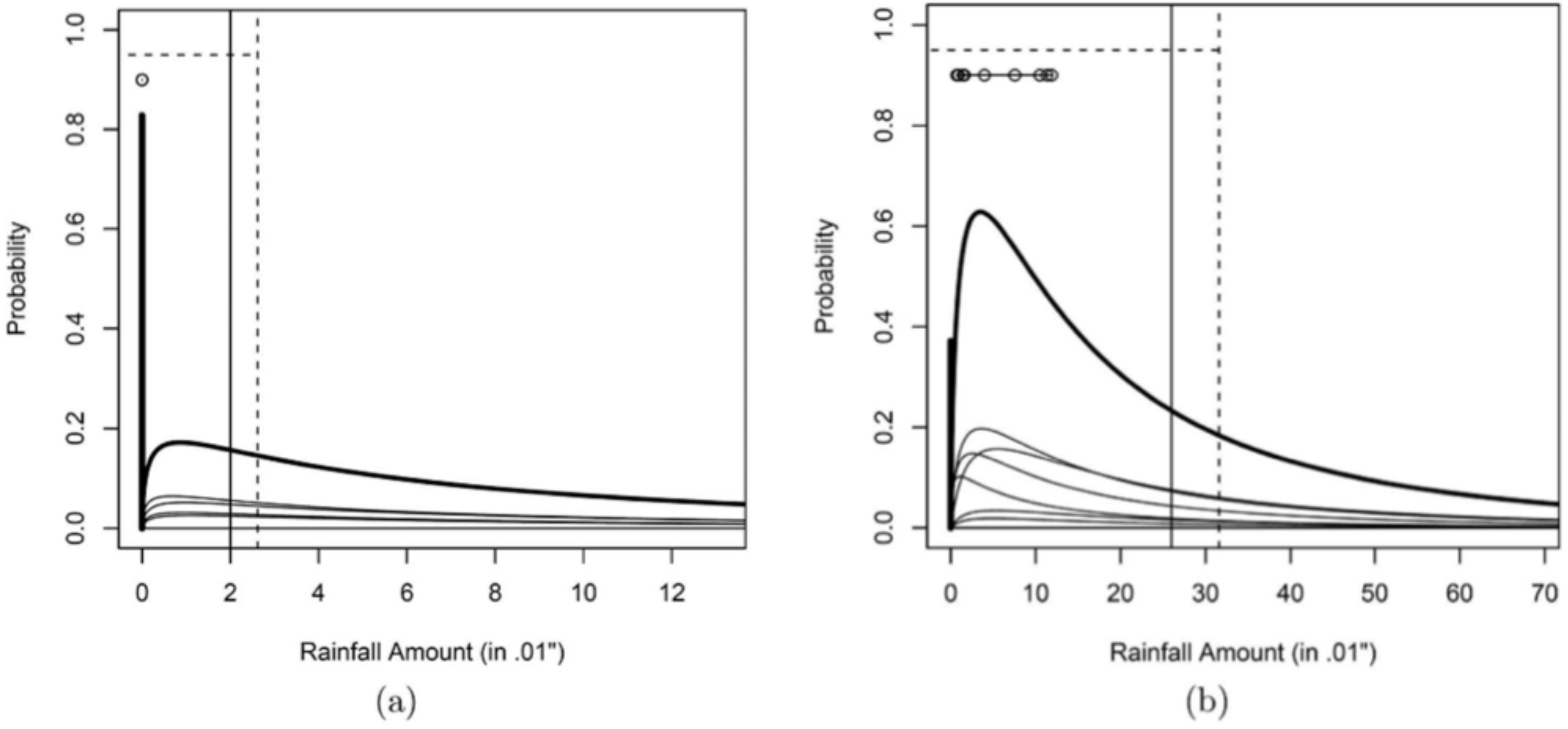


FIG. 3. BMA-fitted PDFs for (a) station KCLM on 19 May 2003and (b) station KPWT on 26 Jan 2003.

Illustration of method shown in Table. 1

Ensemble member	CENT	AVN	CMCG	ETA	GASP	JMA	NGPS	TCWB	UKMO
Station KCLM on 19 May 2003									
BMA weight	0.00	0.39	0.00	0.30	0.19	0.13	0.00	0.00	0.00
Member PoP	0.19	0.16	0.21	0.18	0.17	0.21	0.22	0.23	0.19
BMA PoP	0.17								
Member forecast	0	0	0	0	0	0	0	0	0
BMA forecast	0								
BMA upper bound	3								
Observation	2								
Station KPWT on 26 Jan 2003									
BMA weight	0.00	0.30	0.23	0.23	0.00	0.00	0.03	0.16	0.05
Member PoP	0.46	0.64	0.74	0.59	0.44	0.44	0.70	0.46	0.72
BMA PoP	0.63								
Member forecast	2	8	10	4	1	1	11	1	12
BMA forecast	3								
BMA upper bound	32								
Observation	26								

Table 1. Raw ensemble, logistic regression PoP, and BMA forecasts for two example stations.

Fig 4 shows the model output in the study area

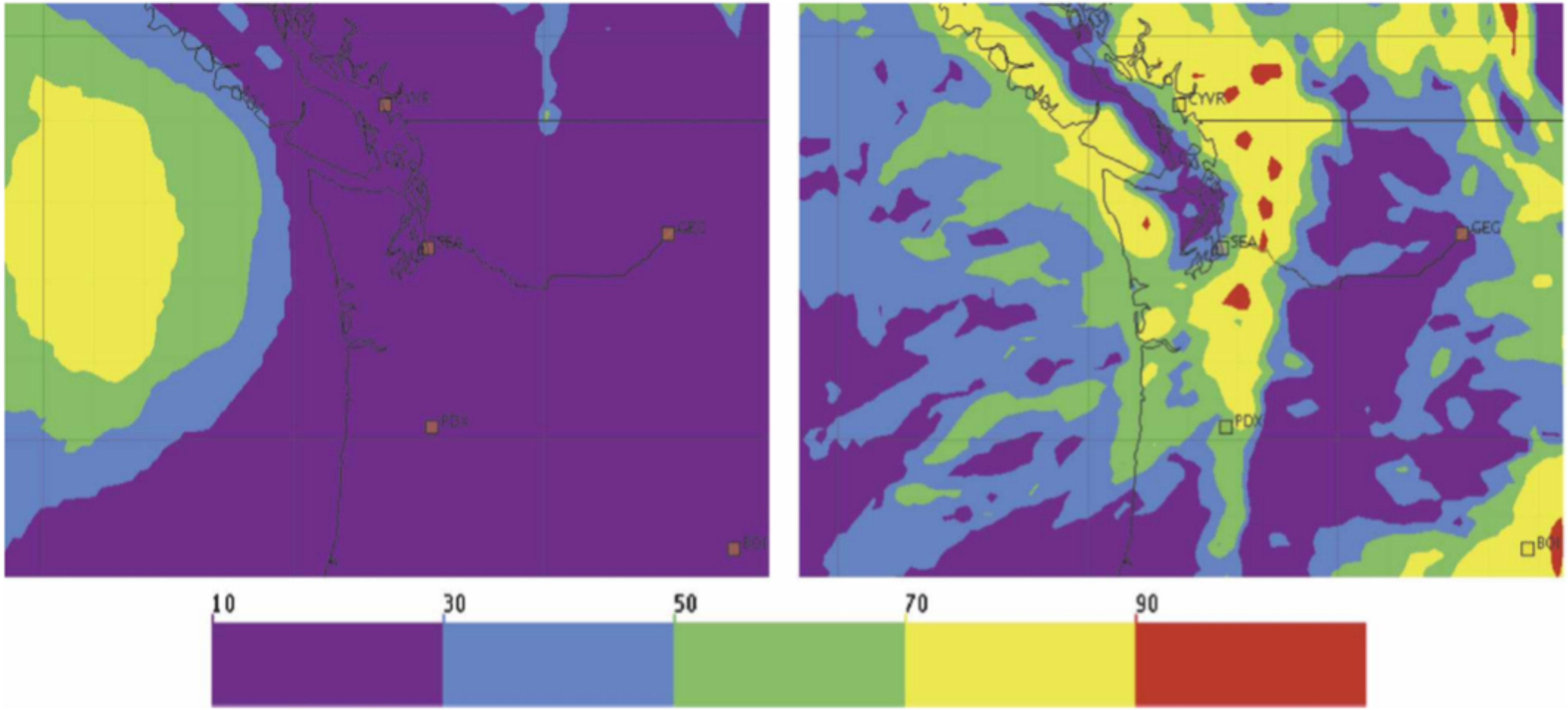


FIG. 4. BMA PoP forecast for (left) 19 May 2003 and (right) 26 Jan 2003.

Fig 5 shows the model output

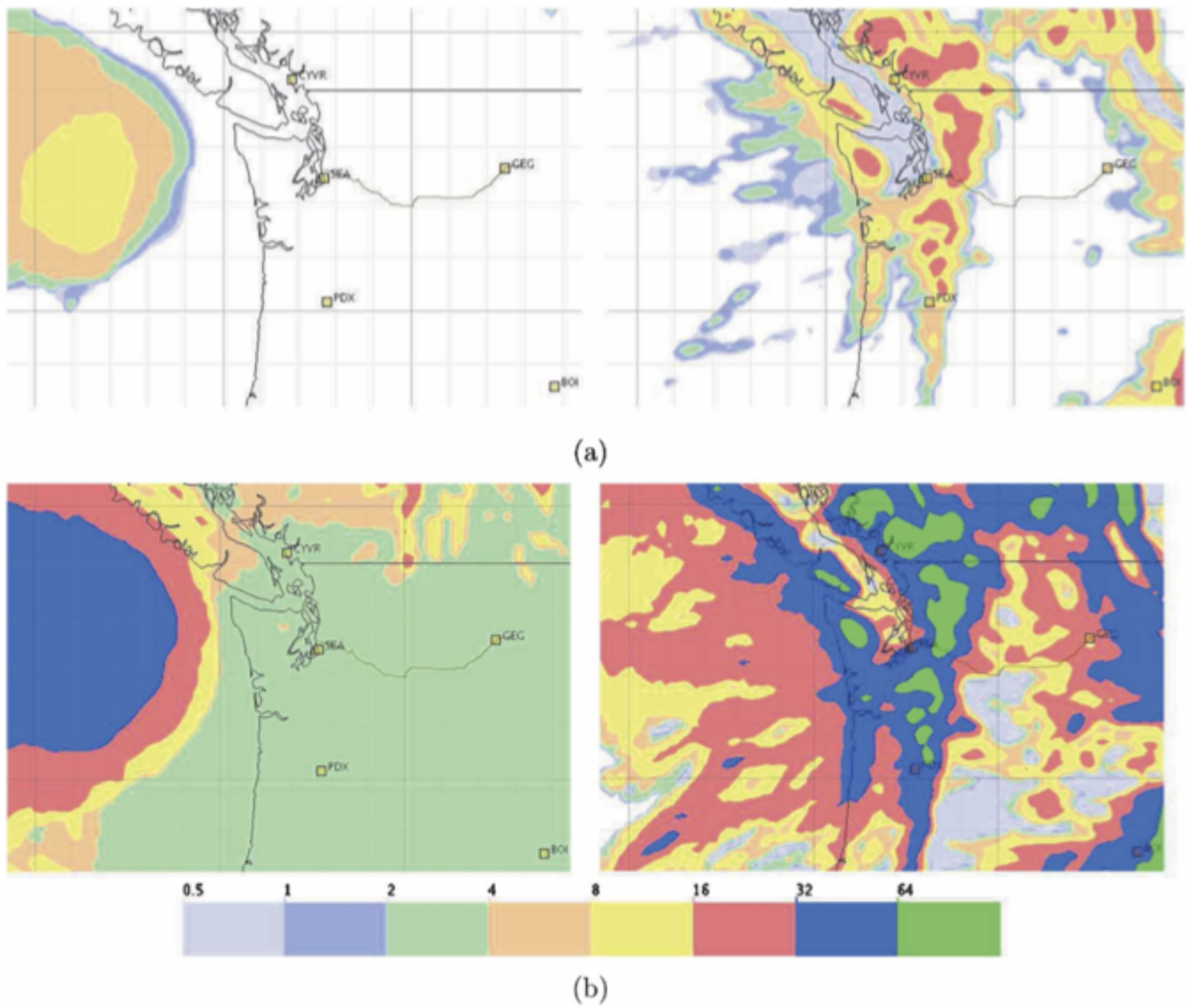
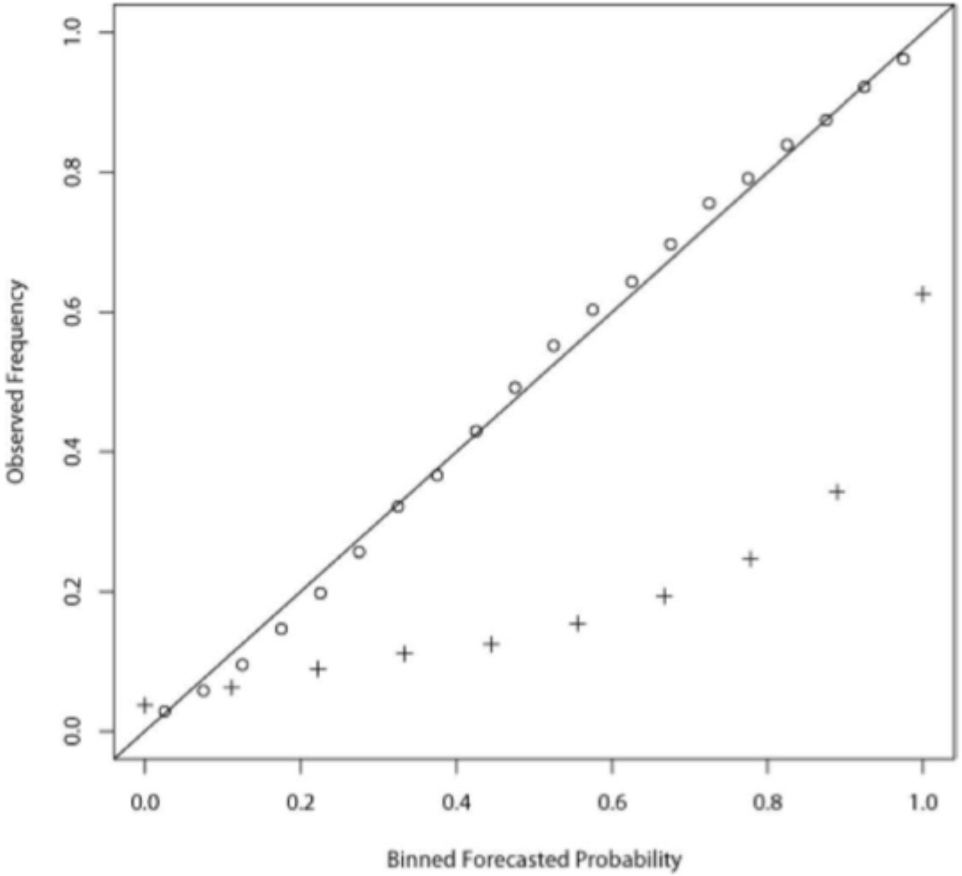


FIG. 5. (a) BMA deterministic forecast and (b) BMA 90th percentile upper bound forecast for (left) 19 May 2003 and (right) 26 Jan 2003, in hundredths of an inch.

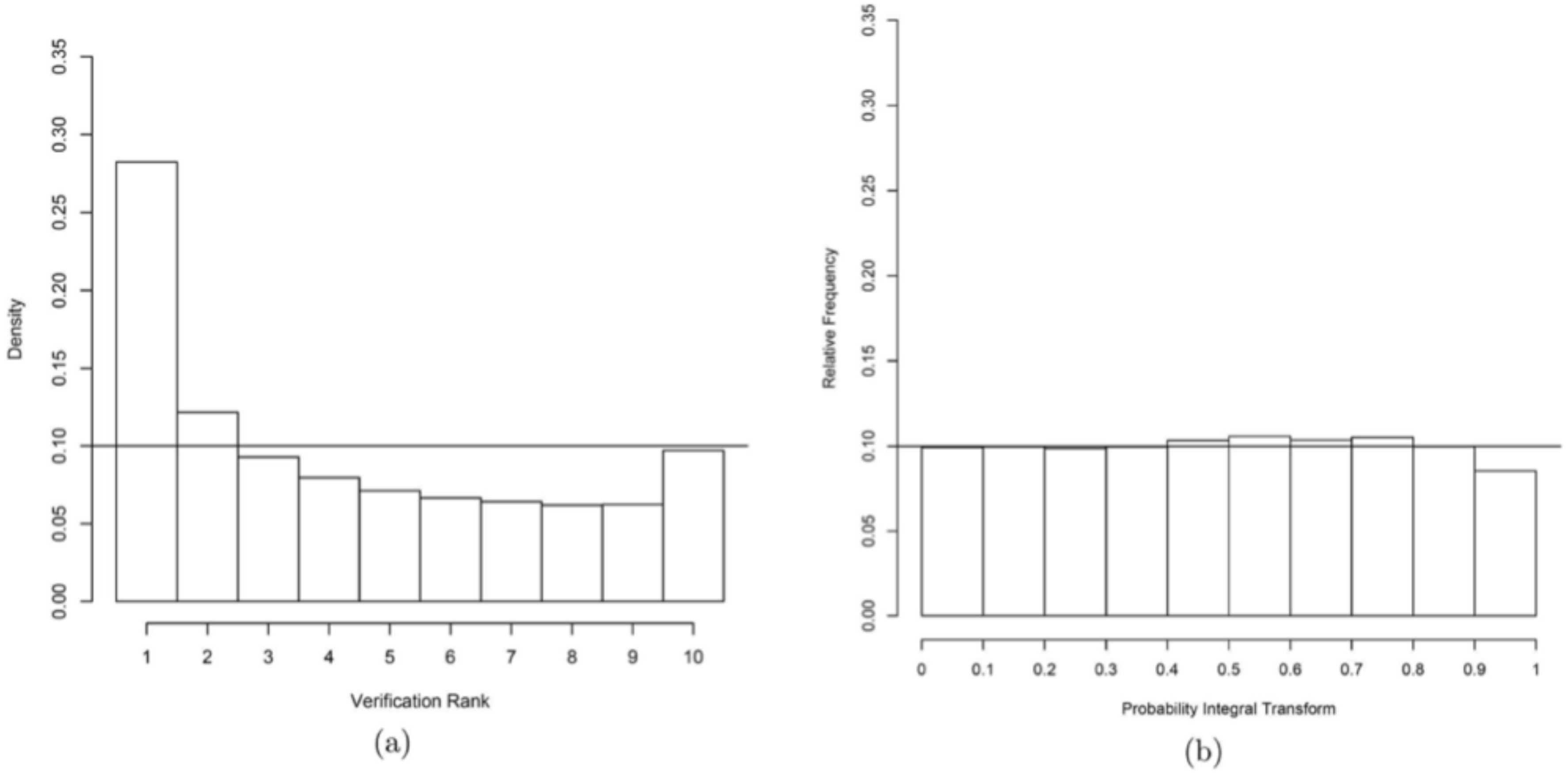
Fig. 6 shows BMA model is reliable



BMA produced well-calibrated results, while the raw ensemble produced severely uncalibrated results.

FIG. 7. Reliability diagram of binned PoP forecast vs observed relative frequency of precipitation, for consensus voting of the raw ensemble (crosses) and BMA (circles).

Fig. 7 shows BMA model is well calibrated



BMA forecast distributions were considerably better calibrated than the raw ensemble.

FIG. 8. (a) Verification rank histogram for raw ensemble fore-casts and
(b) PIT histogram for BMA forecast distributions of precipitation accumulation.

Table 2 shows BMA forecast is more accurate

Score	Threshold	Sample climatology	Ensemble forecast	BMA forecast
MAE		8.7	9.4	7.5
CRPS		7.8	7.6	5.6
BSS	0		−0.18	0.38
BSS	5		0.00	0.36
BSS	10		−0.02	0.34
BSS	25		−0.02	0.31
BSS	50		−0.02	0.26
BSS	100		0.05	0.21
BSS	150		0.05	0.17
BSS	200		0.11	0.14
BSS	250		0.10	0.11
BSS	300		0.09	0.09
BSS	350		0.10	0.08
BSS	400		0.05	0.07

BMA is accurate at both low and high threshold, while raw ensembles lose forecast fidelity at low threshold

Table 3 shows the BMA model is calibrated and sharp

TABLE 3. Coverage and average width of lower 50% and 90% prediction intervals for precipitation accumulation, in percentages and hundredths of an inch, respectively.

	Coverage		Width	
Interval	50%	90%	50%	90%
Sample climatology	50.0	90.0	0.0	24.0
Ensemble forecast	68.5	92.9	11.8	24.2
BMA forecast	50.7	91.1	3.2	22.8

The BMA intervals were close to being calibrated. The BMA PDFs produced narrower intervals than the raw ensemble forecasts for both intervals considered。

ABSTRACT

Bayesian model averaging (BMA) is a statistical way of postprocessing forecast ensembles to create predictive probability density functions (PDFs) for weather quantities. It represents the predictive PDF as a weighted average of PDFs centered on the individual bias-corrected forecasts, where the weights are posterior probabilities of the models generating the forecasts and reflect the forecasts' relative contributions to predictive skill over a training period. It was developed initially for quantities whose PDFs can be approximated by normal distributions, such as temperature and sea level pressure. BMA does not apply in its original form to precipitation, because the predictive PDF of precipitation is nonnormal in two major ways: it has a positive probability of being equal to zero, and it is skewed. In this study BMA is extended to probabilistic quantitative precipitation forecasting. The predictive PDF corresponding to one ensemble member is a mixture of a discrete component at zero and a gamma distribution. Unlike methods that predict the probability of exceeding a threshold, BMA gives a full probability distribution for future precipitation. The method was applied to daily 48-h forecasts of 24-h accumulated precipitation in the North American Pacific Northwest in 2003–04 using the University of Washington mesoscale ensemble. It yielded predictive distributions that were calibrated and sharp. It also gave probability of precipitation forecasts that were much better calibrated than those based on consensus voting of the ensemble members. It gave better estimates of the probability of high-precipitation events than logistic regression on the cube root of the ensemble mean.

Table 1

Figure 1

Figure 3

Table 3

Figure 4&5

Figure 7, table 3

Table 2